APPLYING FREELY AVAILABLE REMOTE SENSING DATA PRODUCTS TO IMPROVE NATURAL RESOURCE MANAGEMENT: CASE STUDIES OF STREET TREE BENEFITS ANALYSIS AND WETLANDS DETECTION

A THESIS

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE

MASTER OF SCIENCE

BY

DAN LANGE

DR. ADAM BERLAND – ADVISOR

BALL STATE UNIVERSITY

MUNCIE, INDIANA

DECEMBER 2017

TABLE OF CONTENTS

Continued

CHAPTER 3: EXPLORATIONS IN WETLANDS DETECTION USING FREELY AVAILABLE DATA PRODUCTS

ACKNOWLEDGEMENTS

Firstly, I would like to sincerely thank my graduate advisor, Dr. Adam Berland, for his exceptional guidance throughout my graduate career, for pushing me to excel, and for continuing to be my most valuable resource as I complete this thesis. I could not have done it without him.

I would also like to thank Dr. Donald Ruch and John Taylor, for their invaluable knowledge based on years of experience with Indiana wetlands, and for their continued support in my academic endeavors.

I would like to acknowledge RCAS members Jon Creek and Jeff Ray, for their patience and support as I worked through this project, and for the opportunity which they have given me.

I would also like to extend a thanks to Mike Planton, who has helped to open new pathways in my life, for which I am grateful.

Finally, I would like to recognize my friends and family for their unending support and encouragement in this endeavor and in life. Thank you so very much.

LIST OF TABLES AND FIGURES

TABLES

FIGURES

EQUATIONS

CH. 1. REVIEW OF URBAN TREE BENEFITS AND GOOGLE STREET VIEW AS A REMOTE SENSING TOOL

Abstract

Urban greenery is an important feature of metropolitan landscapes which provides a number of benefits including increased property values, improved air quality, benefits to human health, mitigation of stormwater runoff, and carbon reduction. Street trees, or trees which occur in municipally owned medians and planting strips along streets, are an integral part of the urban forest because of their manageability at the municipal level, their close proximity to impervious surfaces, and their high public visibility. This review describes how Information about the benefits provided by these trees can be useful for making informed management decisions at the municipal level. However, determining these benefits traditionally requires extensive fieldwork which is often costly and time consuming. Remote sensing methods are available which reduce the need for field work, but may use costly data products or highly technical methodologies. Many local governments or community groups may not have the resources to conduct a traditional field or remote survey; however, it may be possible use only freely available software and data products to provide a baseline understanding of street tree benefits. For example, I-Tree Streets, produced by the United States Forest Service, is a freely available program which calculates benefit values for street tree surveys, and Google Street View, which is also publicly available, provides street level panoramic imagery across the United States and internationally. As discussed in this chapter, these two products in conjunction may be able to be used to conduct a virtual survey mimicking traditional field practices which can be used to calculate street tree benefits with reduced time and monetary investment.

1. Introduction

Municipally managed trees are an increasingly important resource in metropolitan areas and provide a number of ecological, social, and economic benefits including stormwater capture, urban heat island mitigation, improved air and water quality, and increased property values (McPherson et al. 2005; Nowak et al. 2013c). Municipal trees are those trees which are planted, maintained, and managed by local or city government entities. This includes trees within public parks and gardens, and street trees, which are those trees planted in the easements along local streets. Municipal trees are an important component of the urban forest, which includes all public and private trees in a city. Street trees are of particular interest because they are planted and managed by city government entities and impact the way the public perceives their municipal environment more than trees on privately owned residential lots. A municipally planted street or park tree can provide annual benefits of up to 300% of the initial investment (McPherson et al. 2005). Due to its critical role in public infrastructure and the numerous benefits trees provide, the urban forest has been the subject of many research inquiries over the years and continues to be of interest for urban forest management. Below I review the literature on urban forest benefits, with particular emphasis on municipal street trees. Then I describe technological products that can advance our understanding of the urban forest.

2. Benefits of Urban Trees

2.1. *Property Value*

Urban forestation is an integral feature of residential areas, and provides an improved sense of place and community image for municipal areas with urban tree plantings. As an enhancement to the urban environment, the presence of trees translates to increased property values for a home and its neighbors (Escobedo et al. 2014). To assess the scale of this effect, Donovan and Butry (2010) used a hedonic price model to determine if street tree incidence affected sale price and time on the market for home sales in

Portland, Oregon. Houses sold between July 2006 and April 2007 were surveyed, and data was collected about the property location and the street trees on site growing in the public easement or nearby medians. Their analysis showed that the number of trees and size of their crowns for street trees within 100 ft of the home added on average 3% to the sale price of the home (around $8,870) and reduced the time on the market by 1.7 days (Donovan and Butry 2010). Similarly, Sander et al. (2010) used a hedonic price model to quantify property value increase from urban tree cover within 100 and 250 meters of homes in Dakota and Ramsey Counties, MN. At 100 m a 10% tree cover increase boosted the property value by $1,371, and at 250 m, $836 was added to the home value.

2.2. *Air Quality*

The production of oxygen by trees is common knowledge; however, research shows that the oxygen production by urban trees is not consequential enough to be considered a primary benefit. In the United States, urban trees are estimated to produce around 61 million metric tons or oxygen annually, but this amount is considered negligible when compared to the amount of oxygen already in the atmosphere (Nowak and Hoehn 2007). What is of greater importance is the removal of pollutants in the air by urban vegetation, which improves air quality. Trees can remove a number of pollutants from the air including gases (carbon dioxide, ozone, and carbon monoxide) and particulate matter. Nowak et al. (2006a) estimate that pollution removal by urban trees in the United States provides a value of $3.8 billion annually by removing 711,000 metric tons of pollutants. Research also suggests that areas with a higher prevalence of street trees may have fewer instances of respiratory diseases such as early childhood asthma (Lovasi et al. 2008).

Airborne particulates are removed from the air by trees when they are deposited onto vegetative surfaces, with rates varying according to wind speed and precipitation levels (Nowak et al. 2013b). Of particular concern in relation to urban trees are those airborne particulates less than 10

microns (PM_{10}) and those less than 2.5 microns ($PM_{2.5}$). $PM_{2.5}$ is a pollutant regulated by the U.S. Environmental Protection Agency (EPA) that can cause significant health risks related to lung and heart function. Nowak et al. (2013) investigated the $PM_{2.5}$ removal rate of urban trees using EPA monitored $PM_{2.5}$ concentration levels, daily leaf surface area estimated using the i-Tree Eco model, wind speed, and precipitation data for 10 cities in the United States. Results showed removal of $PM_{2.5}$ varies largely based on location with values ranging from 4.7 to 64.5 metric tons removed annually, reducing mortality incidences and saving 1 to 8 lives per year. The removal of $PM_{2.5}$ from the atmosphere amounts in an estimated average health benefit of $1,600 per hectare of tree cover. $PM_{2.5}$ pollution reduction by urban trees results in larger health improvements than PM_{10} reduction, however more PM_{10} can be removed from the atmosphere by urban trees (Nowak et al. 2013b).

2.3 *Human Health and Public Perception*

Along with the removal of harmful pollutants from the atmosphere, urban vegetation provides additional social and physical benefits to human health. Exposure to natural environments promotes lower diastolic blood pressure, higher parasympathetic nervous activity, and lower sympathetic nervous activity, biological responses which are linked to greater levels of relaxation (Tsunetsugu et al. 2013). In addition, urban green spaces may encourage residents to spend more time outdoors, stimulate physical activity, help to alleviate stress, anxiety, and depression, increase personal sense of safety, and reduce crime rates (Kondo et al. 2015; Tzoulas et al. 2007). Kardan et al. (2015) examined the relationship between street trees and public perception of health. This study analyzed satellite imagery and tree data and compared it to information about public health from the Ontario Health Study, finding that residents in areas with higher levels of urban forestation have better health perception and experience fewer cardio metabolic conditions at a rate equal to a $10,000 income increase for city blocks with more than 10 trees.

Street trees have a large impact on the visual attractiveness of urban residential areas (Schroeder and Cannon 1983). Most of the public views urban vegetation positively and acknowledges the value of the benefits provided, most often referring to the aesthetics, shade, and the property value increase associated with street tree plantings (Mullaney et al. 2015). Post hurricane Hugo, 30% of the residents of Charleston, South Carolina, interviewed via telephone by Hull (1992) identified the urban forest as the most significant feature damaged, citing the urban forest's contribution to community image, environmental quality including energy benefits, and the sense of place provided. Commercially, urban greenery can have a positive psychological impact on consumers and may increase revenue district-wide in sectors with urban forestry programs (Wolf 2003).

2.4 *Mitigation of Stormwater Runoff*

Urban trees mitigate a portion of stormwater runoff by interrupting the path of precipitation. Reducing stormwater runoff helps to enhance downstream water quality, as less debris (including chemicals from lawns, salts and deicers, and particulates from the built environment) is washed downstream. One way that trees help to mitigate stormwater is through their leaves and branches, which delay stormwater flow during rainfall events, allow for some evaporation to occur, and can redirect water flow down the trunk of the tree and into the soil (Bartens et al. 2008; Mullaney et al. 2015). Furthermore, soil infiltration is increased in tree root zones as roots are able to penetrate compact soils and provide avenues for water flow, additionally decreasing total runoff by increasing soil infiltration in compacted soils 62 - 153% (Bartens et al. 2008). In addition, urban vegetation can reduce soil erosion by creating a barrier between rainfall and bare surfaces (Xiao et al. 1998). There is also evidence that transpiration by urban trees substantially increases the capacity of the soil to infiltrate stormwater, because of the water cycling from the soil to the atmosphere. More research is needed to accurately quantify this process (Berland et al. 2017).

Rainfall interception describes the process by which urban trees detain incoming precipitation on leaf and bark surfaces, preventing the precipitation from contributing to runoff. Annually, an urban forest can reduce runoff by 2 – 7%, but when also considering the landscaping structures in which trees commonly grow, up to 65% of residential storm runoff can be reduced (Fazio 2010). Xiao et al. (2000) found that per tree rainfall interception was on average 15% per rainfall event for precipitation falling in the vicinity of a 9 year old pear tree (*Pyrus calleryana),* a very common street tree species, with interception more effective in shorter and lighter rainfall events. Further studies have estimated annual rainfall interception by tree-covered areas in Sacramento County, California at 11.1%; and, for the street and park trees in Santa Monica, California, 1.6% of the total precipitation was intercepted, providing a value of $110,890 annually (Xiao et al. 1998; Xiao and McPherson 2002). More recently, Berland and Hopton (2014) measured street trees on 597 randomly selected street segments in nine Cincinnati neighborhoods using the i-Tree Streets model and found an estimated average stormwater interception of 120.4m^3 per km of street length at a value of $179.45 per km. Estimated interception values may seem relatively low because they only account for water retention by the stems and leaves of the plants. When taking account for the soil beneath the trees in medians, cutouts, or planting strips, stormwater may be reduced by up to 62%, as a large portion of the precipitation infiltrates into the soil around a tree (Armson et al. 2013).

2.5 *Temperature Regulation*

Urban trees reduce building energy costs by casting shade in the summer, and blocking wind in the winter. Casting shade reduces the amount of direct sunlight on buildings and impervious surfaces which prevents the surfaces from absorbing as much solar radiation and reduces the need for air conditioning. Blocking the wind prevents warm building surfaces from being cooled by winter winds and reduces heating costs. Annually, a single street tree can reduce energy costs by $2.16 to $64.00 which averages

at $15.00 or 95kWh per year, depending on various factors including the tree's size, species, and location (Mullaney et al. 2015). In Sacramento County, California, Ko et al. (2015) found that 20 – 22 year old trees provided an average annual energy savings of 80kWh per tree. It is important to note that these functions work in reverse as well, shade from evergreen trees in the winter can potentially raise heating costs, and wind barriers in the summer can potentially trap heat against building surfaces or reduce wind infiltration into interior spaces; however, net benefits remain positive (Wang et al. 2014).

Additionally, street trees can reduce outdoor temperatures, cooling urban forested areas between 5°C and 20°C, and reducing the urban heat island effect (Mullaney et al. 2015). Urban heat island effect is the term used to describe the higher average air temperature and lower wind speed for urban areas caused by the large concentration of built-up surfaces that absorb solar radiation (Bolund and Hunhammar 1999). This effect can increase urban air temperature by 5-8°C when compared to surrounding countryside, leading to more energy use for temperature control, more water use for landscaping, and a higher frequency of public health problems (Hardin and Jensen 2007). Millward et al. (2014) quantified the effect of urban vegetation on summer building surface temperatures in Toronto, Canada. The study used multiple pairs of temperature loggers, one of each pair placed on a building surface in full sunlight, and one shaded by vegetation. Building surface temperatures averaged 11.7°C lower in areas shaded by vegetation, with results varying by vegetation type, size, and location relative to building surfaces. Another study by Hardin and Jenson (2007) used leaf area index (LAI) and thermal imagery to quantify the inverse relationship between urban forest density and summertime surface temperatures for Terre Haute, Indiana. LAI is an estimation of the amount of foliage intercepting light and interacting with atmospheric gases. Linear regression analysis showed that leaf area accounted for 62% of surface temperature variation, with increasing leaf surface area correlated to decreasing surface kinetic temperatures.

2.6 *Carbon Reduction*

Trees take in carbon as they grow, and carbon which is utilized for tree growth is no longer an active greenhouse gas, which helps to mitigate climate change. Carbon which is contained within the branches, trunk, and root systems of trees long term is referred to as “carbon storage,” whereas carbon which is absorbed during a single season to create new growth is termed “carbon sequestration.” Both carbon storage and carbon sequestration reduce atmospheric greenhouse gas levels. Carbon emissions are further avoided by the effect on heating and cooling that urban forestation can provide (McPherson et al. 2013). Using less energy for temperature control means that less carbon is put into the atmosphere through coal burning power plants and natural gas consumption. For urban areas in the United States, Nowak et al. (2013a) estimated total carbon storage in the United States at 643 million metric tons or a value of $50.5 billion in avoided costs, and annual sequestration at 25.6 million metric tons or a $2.0 billion value nationally for 2005. Soils in urban areas are highly likely to store more carbon than the vegetation; however, because of the ephemeral nature of urban forests the carbon within them is of greater concern. Of note, however, is that many tree management practices, planting and removal, pruning, and pest treatments all produce carbon emissions with equipment they require, thereby negating some or all of the carbon reduced by the tree’s growth. In addition, the carbon sequestered by a tree over its lifetime has the potential to be re-released into the atmosphere as the tree decomposes.

2.7 *Urban Tree Disservices*

Urban trees provide a great number of benefits, but their potential disservices should also be taken into consideration. Urban trees can pose a potential financial burden to both cities and homeowners when trees become diseased, or an invasive pest spreads throughout the area. The emerald ash borer, a more recent invasive pest which has spread at alarming rates, is estimated to cost municipalities and home owners more than $1 billion annually (Siegert et al. 2014). Municipally, the cost of tree management can

add up quickly when considering the pruning necessary to keep trees away from powerlines and out of street corridors, the damage to sidewalks, roads, and buildings caused by root infiltration, and management of tree-borne pests. In addition, some management practices, including fertilizer and pesticide application, may offset the benefits urban forestation provides to stormwater quality; and, as previously mentioned, management involving equipment may partially negate urban tree carbon reduction (Escobedo et al. 2011). Urban trees also provide shelter and sustenance for insects and wild animals, which can potentially spread disease and cause damage to landscaping and buildings. In addition, improper species choice may lead to the spread of invasive plant species or increase pollen related allergens throughout an area. Proper planting location and species choice (e.g. tall growing trees should not be planted under powerlines) is therefore of great concern as it can potentially eliminate many of the maintenance costs and alleviate potential disservices of urban tree plantings.

3. Street Trees are an Integral Part of the Urban Forest

Street trees are those trees planted within the right-of-way or easement along public streets. This includes plantings in medians, sidewalk cutouts, and in planting strips, the grassy area between the street and the sidewalk. These trees account for around 10 – 20% of the total number of trees in any urban forest; however, street trees are a critically important part of the urban forest because they are the most abundant and widespread municipally managed trees (McPherson and Rowntree 1989). The remainder of the urban forest occurs largely within private residential lots, which are managed privately by the homeowner. As previously discussed, urban forestation provides many environmental, social, and economic benefits. Street trees are a large contributor to these benefits due to their proximity to impervious surfaces and vehicular traffic, and their high visibility within municipal and residential areas. Thus, it is important to be able to quantify the street tree population in terms of numbers, sizes,

diversity of species, and locations, and to have the capability of determining the benefits provided in order for quality management to occur at the municipal level.

The structure, diversity, and benefits provided by street tree populations have been the focus of many research inquires (Armson et al. 2013; Cowett 2012; Denman 2006; Fischer and Steed 2008; Gorman 2004; Landry and Chakraborty 2009; Livesley et al. 2014; Lovasi et al. 2008; Sanders 1981; Schroeder et al. 2006; Soares et al. 2011; Stovin et al. 2008). McPherson et al. (1997) quantified the benefits of Chicago's urban forest population using canopy cover analysis and *in situ* data collection methods. In 1991, Chicago's street tree population accounted for approximately 10% of the total trees in the city and provided an estimated value of $9.2 million in air quality improvements, with a net value for all services measured of $402 per tree planted. In 2007, the street trees in the city of Portland, Oregon, were estimated to boost property values for homes in the city by $1.35 billion (Donovan and Butry 2010). Kovacs et al. (2013) estimated the potential value provided by carbon reduction for potential street tree plantings based on available planting space in New York City, New York, and found that in the next 100 years, $3,133 - $8,888 per metric ton of carbon could be saved, with differences based on species planted.

4. Data Collection for Urban Forest Research

Many research studies interested in quantifying urban forest attributes depend on *in situ* data collection methods. Field surveys can often be very costly, as they usually require many hours of work from trained professionals, and become more expensive if they supplement with hyperspectral or LiDAR remote sensing products (Wegner et al. 2016). As an example, it took five months for a six person team to conduct a 5,600 tree inventory in Ithaca, New York, and 14 weeks for a two person team to survey 200 circular 0.04 ha sample plots at Auburn University in Auburn, Alabama (Martin et al. 2013). Time and expenses can be reduced by using volunteers or "citizen scientists," but at some risk to data quality

as many laypeople are untrained at species identification and common tree measurement practices (Roman et al. 2017).

Given the challenges of collecting reliable *in situ* data, some studies have relied mainly on remotely sensed data. Remote sensing approaches have the potential to reduce the amount of time the study requires, and as more and more data products become freely available, remotely sensed tree inventories become a more cost effective option. O'Neil-Dunne et al. (2014) used LiDAR, hyperspectral imagery, and thematic GIS layers to map the urban tree canopy for 70 locations in North America using geographic object based image analysis (GEOBIA) and produced high resolution tree canopy map products. The GEOBIA approach has the potential to be used municipally for producing timely tree canopy maps, but relies on knowledgeable GIS professionals to execute the analysis as well as the availability of LiDAR and other data products to keep costs down. Hyperspectral and LiDAR data were also used by Alonzo et al. (2016) to remotely identify tree species and to calculate LAI and carbon storage for the urban forest in downtown Santa Barbara, California. The authors then compared results to plot sampled field data from the same area and found only 5% difference in canopy size estimates, an 11% difference in LAI estimates, and a 9% difference in carbon storage estimates. This study shows that urban forest benefits can in fact be accurately estimated using remote sensing methods, provided appropriate data products are available. Landry and Chakraborty (2009) used a maximum likelihood classification to identify tree cover in satellite imagery for the purpose of comparing street tree canopy cover among census block groups with different socioeconomic conditions. In Berland (2012), tree canopy cover change over time was examined using visual interpretation methods using aerial photography from 1937 to 2009 for Minnesota's Twin Cities. This study used aerial photography only, as a means to estimate tree canopy by generating and classifying random points within the images. Because of the increasing public availability of aerial photography, this method provides easy access to basic urban forest benefits for municipal managers. A similar method is used by i-Tree Canopy

(https://canopy.itreetools.org/), a freely available simple urban forest benefits calculator from the U.S. Forest Service. I-Tree Canopy automatically generates random points over a user defined area and prompts users to visually identify the land cover at the point (lawn, house, tree, etc). Aerial imagery can be used to estimate tree canopy cover, but identifying other aspects of the urban forest requires more information. Data including genus, species, tree health, and diameter at breast height (DBH) are all often used in calculations of urban forest benefits, but are difficult to obtain using traditional remote sensing data products. As described in Section 6 below, visual or automated interpretation of ground based imagery could provide the missing elements for virtual street tree surveys.

5. I-Tree: A Tool for Urban Forest Benefits Analysis

Developed by the United States Forest Service, i-Tree (http://www.itreetools.org/) is a suite of urban forest benefits analysis tools which are freely available to the public. The i-Tree software suite does not require extensive technical knowledge or expertise, so it can be used by any interested parties to understand and quantify the value of their municipal forest. However, conducting a manual i-Tree survey using i-Tree protocols can be both expensive and time consuming.

The flagship Forest Service models were originally released as the Urban Forest Effects model (UFORE) and the Street Tree Resource Assessment Tool for Urban Forest Managers (STRATUM), but are now called i-Tree Eco and i-Tree Streets, respectively (Cowett 2014). I-Tree Eco is based on circular plot sampling or complete inventory methods and can be used to quantify an urban forest, like a large city park or university campus, in its entirety. A 100% i-Tree Eco inventory at Auburn University in Auburn, Alabama, estimated a value of $10 million for the 238 ha campus forest (Martin et al. 2011). I-Tree Streets focuses on street trees by using random samples from street segments throughout a municipal area to estimate tree population characteristics and associated benefits.

I-Tree Streets was released to the public in 2006, and was specifically designed with urban forest management in mind (https://www.itreetools.org/). The model uses data from in-depth studies of "reference cities" to calculate urban tree benefits, with one reference city for each of the sixteen national climate zones in the United States. Reference cites were selected for their pre-existing tree inventory data, tree species and size diversity, and available planting records. Trees of each major species within the reference cities were measured for DBH, crown volume, and leaf area, to create predictive models for user input (McPherson 2010). The i-Tree model also utilizes weather and pollution data to tailor tree benefits to the specific area of interest. I-Tree users can calculate urban forest benefits based on a minimum of just two measurements, DBH and species. Using the data from the reference cities, these measurements are correlated to tree canopy size and shape and are used for the calculation of benefits like air pollution removal for different pollutants, stormwater interception, and increased property values.

Both i-Tree Streets and Eco have been utilized and evaluated by a number of research studies (e.g., Baró and others 2014; Baumgardner and others 2012; Cowett 2012; Martin and others 2011; McPherson and Kotow 2013; Millward and Sabir 2011; Nowak and others 2013c; Nowak and others 2006b; Soares and others 2011) . At Allan Gardens, a municipal park in Toronto, Canada, STRATUM was used to estimate the park's benefits. Results showed that each of the park's trees provided an average value of $95 annually in environmental and aesthetic benefits to the city (Millward and Sabir 2011). Berland and Hopton (2014) used i-Tree Streets field survey protocols to quantify street tree basal area for multiple neighborhoods with differing Tree City USA status in Cincinnati, Ohio, finding that those communities involved with the program had more street trees. Field data for this study were collected over a 3-month period, producing a sample of 10.8% of the total number of streets within the study area. Soares et al. (2011) used STRATUM to quantify street tree benefits in Lisbon, Portugal. The Lisbon street tree population was sampled using 65 randomly generated street segments and field work was

completed in the summer of 2004. The street tree population was found to provide a net benefit of $6.5 million annually with property value being the largest contributor to the benefits provided. As i-Tree Streets bases its model on sixteen U.S. reference cities, this study, which does not fall into one of the reference climate zones, matched growth curves from Lisbon trees to trees in multiple reference cities to improve final results.

6. Google Street View as a Source of Remotely Sensed Imagery

Google Street View (GSV) (https://www.google.com/streetview/) is a freely available ground based imagery service. Officially launched in 2007, GSV began as a collaborative research project with Stanford University (Wu et al. 2014). This initial research focused on multiperspective image compilation of streetscapes using imagery extracted and stitched together from video streams (Roman et al. 2004; Roman and Lensch 2006). Since then, GSV has developed into a massive venture, operating in over 20 countries around the world and providing nearly complete coverage in US cities. GSV provides geographically referenced panoramic imagery, mostly taken along road networks, collected using car mounted sensors (although sometimes mounted on bicycle, backpack, or snowmobile, as needed). The data collection technology has seen many renditions, but commonly features cameras, laser rangefinders and GPS data collection. A large portion of the imagery is now collected with custom panoramic camera systems that Google produces in house. These systems, depending on iteration, feature 8 or 15 5-megapixel CMOS image sensors, instead of traditional shutter cameras, and feature no moving parts (Anguelov et al. 2010).

As a potential remote sensing tool for research, GSV has been largely underutilized, but a small number of studies have evaluated the potential of the imagery service. One application of GSV imagery is as a tool for conducting virtual neighborhood or streetscape audits. Badland et al. (2010) used GSV to classify walking and cycling amenities for street segments in Auckland, New Zealand, and compared

results to a physical audit in the same location. Results showed the virtual audit was completed more quickly and at lower cost while maintaining acceptable agreement with the physical audit. Similar results were found by Clarke et al. (2010) whose virtual neighborhood audit of land use, and recreational and food amenities for a neighborhood in Chicago, Illinois, was completed at a lower cost and less obtrusively than a physical audit for the same location. Wu et al. (2014) conducted virtual and physical audits for postcodes in Cambridgeshire, England, evaluating features related to mental health. This study found the virtual audit to require fewer resources and still provide adequate reliability. One shortcoming of GSV as a virtual audit tool is the imagery date, which may not include recent improvements to the built environment, is not updated on a predictable basis, and may contain disparities in imagery dates between street segments.

While many existing GSV studies relied on manual image interpretation, automatically extracting data from GSV images is possible but requires technical knowledge. Tsai and Chang (2013) outline a method for extracting three-dimensional positions from GSV imagery by comparing the location of objects in multiple adjacent panoramas. In another study, GSV was used to extract the geographically referenced positions of traffic signs for engineering purposes (Yan et al. 2013). The authors used direct linear transformation, a photogrammetry technique, to convert the positions of detected traffic signs into georeferenced locations. Although highly technical, this method could be used to update traffic sign databases automatically, with greater ease than manual surveys.

In relation to urban forestry, GSV imagery has several potential applications. Street trees, which as previously mentioned are of great interest to municipal forestry management, will inherently be visible in GSV imagery as the panoramas are primarily captured along roadways. One GSV based method for quantifying street level greenery was explored by Li et al. (2015b). By requesting parts of a GSV panorama from the GSV application program interface (API) using a custom URL format, one can obtain imagery from a specific location, facing a desired direction. Li et al. (2015b) obtained images from GSV in

this fashion, and then extracted vegetation from the images based on the spectral reflectance of the vegetation in the standard red, green, and blue (RGB) photo bands. The amount of green pixels representing vegetation was then used to calculate a "Green View" index, by comparing the green area of the image to total area of the image. Li et al. (2015a) used this method to examine differences in street level vegetation along socioeconomic gradients in Hartford, Connecticut, finding that areas with higher per capita income levels often have more urban vegetation. A couple of other studies have also employed this methodology; Richards and Edwards (2017) extracted the green pixels from upwards facing GSV API images to quantify canopy coverage in Singapore, finding that vegetation intercepts a median of 8% of the total incoming solar radiation for all street locations. Long and Liu (2017) calculated a Green View Index for over one million street level images in 245 cities using imagery from Tencent Street View, allowing them to compare street level greenery between many Chinese cities. The method used by Li et al. (2015 a,b) to quantify street level greenery can be done autonomously, uses freely available imagery, and does not require any field surveying. However, using GSV imagery in this way has limitations, as it only contains RGB bands so green objects such as road signs are often confused for vegetation. Advanced computer vision techniques are emerging to overcome this barrier by considering shapes and texture in the imagery in addition to color (Seiferling et al. 2017), but these approaches are not widely available and require considerable computing expertise.

Another method of using GSV imagery to conduct an analysis of an urban greenspace would be to manually conduct a virtual survey using visual image interpretation, similar to the aforementioned virtual neighborhood audits. Rousselet et al. (2013) conducted a virtual survey using GSV imagery to examine distribution of a tree pest, in France. In this study the virtual survey was conducted by visually interpreting GSV sample areas for signs of the pine processionary moth, which builds silk nests in pine trees (mostly *Pinus nigra*). This virtually survey method was analogous to field techniques for detecting the moth's presence from roadsides, which was also conducted for comparison. The virtual survey

produced results similar to the field survey at a small scale, but was less accurate at a larger regional scale. A virtual survey conducted in this manner has the benefit that no intense technical work or photogrammetry is required for the completion of the work; instead, it mirrors common field practices.

7. Challenges, Underexplored Topics, Research Priorities

A number of challenges exist for benefits quantification of the urban forest at the municipal level. Many urban forestry programs receive little funding and may not be able to spare the time or money required to conduct an in-depth field survey. Field surveys require transportation to field locations and wages for the survey analysts. Using i-Tree Eco protocols can take around 14 weeks for an experienced two person team to complete the minimum recommended 200 sample plots (Alonzo et al. 2016). The availability of accurate survey analysts/professionals is often scarce, and cities may have to rely on less qualified students or citizen scientists for species identification and tree measurements. This potentially impacts the data quality, leading to erroneous results that can negatively affect management decisions based on field data and subsequent benefits analysis.

GSV imagery also has its limitations. Image quality can vary dramatically between panoramas, with some panoramas containing stitching issues and others containing large obstructions including semi-truck trailers and censored areas. An additional concern for using GSV imagery as a substitute for *in situ* data collection is the date discrepancy between GSV panoramas. Image dates often vary by months or years, even along a single street segment. This potentially adds temporal error into any studies carried out using GSV as a medium for data collection. Although GSV coverage is remarkably good in the United States and a number of other countries, there are still locations without GSV imagery. Sometimes this just amounts to a missed side street or a new development that has yet to be captured, but can potentially be a large area that GSV has not yet surveyed. Also, depending on the country, privacy laws may prevent extensive GSV coverage. For any study relying on visual image

interpretation by multiple analysts, rater variability becomes a concern, but standardized virtual survey procedures should help to uphold data integrity (Rousselet et al. 2013).

In summary, urban greenery is an important resource in metropolitan areas because of the numerous social, economic, and ecological benefits provided. Street trees do not make up the majority of the urban forest, but they play a critical role due to their proximity to impervious surfaces, their high public visibility, and because they are managed by local or city governments. Responsible management of the urban forest can increase the benefits provided, but requires knowledge of the forest's structure, diversity, and spatial characteristics to make educated management decisions. Research has shown that this can be accomplished in several different ways, including field surveys and remote sensing methods; however, from a municipal management perspective, these practices are not always options due to the time, money, technical knowledge, or data products required. However, it may be possible to use freely available software and data products to conduct a street tree benefits analysis. Virtual estimation of ground level imagery from GSV, in lieu of extensive field work, and benefits estimation through i-Tree models could provide a base understanding of urban forest benefits at a fraction of the time and monetary investment required by traditional survey methods.

CH. 2. VIRTUAL STREET TREE DATA COLLECTION USING GOOGLE STREET VIEW AND I-TREE STREETS SAMPLING PROTOCOL

Abstract

Urban greenery is an important factor in metropolitan areas for providing environmental, social, and economic benefits including stormwater capture, improved sense of community, and increased property values. However, assessing these benefits requires extensive fieldwork that is costly and time consuming. With the growing wealth of publicly accessible aerial and street-level imagery data available online, conducting a virtual analysis of urban tree benefits is becoming increasingly feasible. In this study, Google Street View was used to conduct a virtual survey of street trees in metropolitan Cincinnati, OH, and compare results to an existing field survey from the same location. The USDA's i-Tree Streets model was then used to calculate the environmental benefits provided. This research aims to determine whether a virtual survey can be used to generate estimates of urban tree benefits that are statistically similar to estimates derived from field survey data, and to examine virtual survey techniques to determine best practices for visual estimation of street tree attributes.

The Streets model is driven by tree abundances, species, and sizes, so accurate estimation on these attributes in the virtual survey should yield comparable results between the field survey and virtual survey. In this study, the virtual survey generated results similar to the field survey for genera (89% agreement), species (58% agreement), and diameter classes (67% agreement). Net annual benefits calculated by i-Tree for matched virtual survey trees were on average 6.4% different from field survey values. This investigation shows promise for using freely available street-level imagery to conduct virtual analyses of street trees and associated environmental benefits. This virtual approach could save municipalities time and money by reducing the need for field work.

Portions of this research have been published in Berland and Lange (2017). See Appendix 1 for the author's prepress version of this article.

1. Introduction

As discussed in the previous chapter, urban greenery is an important factor in metropolitan areas for a variety of reasons including reducing stormwater runoff, providing energy savings via shading, improving air quality, and improved aesthetic value (Mullaney et al. 2015). Urban forestation programs are becoming more abundant throughout the country as environmental awareness rises. Geospatial quantification of street tree benefits is one way for municipal foresters to make educated management decisions about their tree plantings; however, many municipalities struggle with quantification of their urban forest benefits because most available methods require extensive field work or highly technical imagery classification techniques.

Freely available data products and software, along with virtual survey methods could reduce the time and monetary commitments to conduct traditional benefits assessments for urban trees. I-Tree Streets (https://www.itreetools.org/) is a well-documented and publicly available benefits modeling software from the USDA forest service which calculates urban forest benefits using standard tree measurements and sampling procedures, but requires extensive field data collection. Google Street View (GSV) has been used for several exploratory virtual surveys, however most have focused on automated image extraction using the Google Maps Application Program Interface (API) (e.g., Li et al. 2015a,b, Wegner et al. 2016), or the use of GSV imagery to identify easily recognizable features such as buildings (Badland et al. 2010; Clarke et al. 2010; Less et al. 2015; Wu et al. 2014). Few have evaluated the use of GSV imagery as a virtual survey tool, and none thus far are known to have attempted virtual estimation of urban forest characteristics as a means to quantify benefits provided. This research evaluates the use of GSV imagery as a remote sensing tool for virtual estimation of street tree characteristics for use in i-Tree Streets. This research addresses the growing need for municipal access to urban forestry data while using only freely available data products and benefits modeling software, implementing a simple approach that is accessible to those without advanced computer skills, and

reducing time and monetary costs associated with field data collection. This survey process could potentially be completed by any interested parties for a baseline understanding of their street tree population. The goals of this research are to understand (1) what level of data quality can be expected from a virtual survey of street trees using GSV, and (2) how well virtual survey data can be used in place of field data to estimate environmental benefits provided by street trees.

2. Conducting the Virtual Survey Using Google Street View

2.1 *Software*

To conduct the virtual survey and compare results to the field data, I used a few different computer programs. GSV was used to visually estimate street tree characteristics for the municipal trees visible within the panoramas. GSV can be accessed freely via downloadable programs (Google Earth and Google Earth Pro), or online through Google Maps (https://www.google.com/maps) and the GSV API (https://developers.google.com/maps/documentation/streetview/). I used Google Earth Pro because it supported GIS shapefile visualization, which allowed me to directly view the randomly selected street segments to be included in the virtual survey within the GSV interface. These segments were visible both in the aerial imagery as well as within the Street View panoramas, and were useful for locating segments and keeping track of their start and end points while working along a road segment.

I used ArcGIS to record my virtual survey data in geospatial format to facilitate comparison of virtual data and field data. To prepare data for benefits modeling in i-Tree Streets, I exported tabular data from ArcGIS to Microsoft Excel and reformatted it using instructions available at https://www.itreetools.org/. I-Tree Streets was then used to model the environmental benefits provided by trees recorded in the virtual survey. Depending on the expertise of the virtual survey analyst, other data entry approaches could be implemented; at a bare minimum, the virtual survey requires access to GSV and the i-Tree Streets program.

2.2 *Methodological Approach to Virtual Data Collection*

For this research, extensive data was needed from an existing street tree inventory to provide a basis for comparison. For a practical application of this methodology, a relatively smaller set of ground truth data would be needed to assess the quality of virtual survey data. Data was obtained from a completed study in Cincinnati, OH (Berland and Hopton 2014). I received GIS data files for the Cincinnati neighborhoods and suburban municipalities surveyed in the original study, including the randomly selected street segments within those communities, and the tree data collected. Additionally, I received i-Tree Project files for the three municipalities which I virtually surveyed. It is important to note that the tree data and i-Tree files were not explored prior to completion of the virtual survey, as this would create an unfair bias in tree identification and size estimation. Below I include the major steps taken to complete the virtual survey with specific interest in comparing virtual survey data to existing field survey data. Similar steps could be used for a practical application of this methodology by excluding those extra steps which I took in order to form a comparative relationship between surveys.

1. The original inventory data was obtained, Google Earth Pro and i-Tree Streets were installed. The boundaries of municipalities, street segments, and survey segments were loaded into a blank ArcMap document.
2. The GIS shapefiles of study segments and community boundaries were exported from ArcMap and loaded into Google Earth Pro.
    a. Importing the shapefiles into Google Earth Pro makes it easy to locate and inventory the correct road segments, although Google's address locator could also be used identify the correct survey locations if GIS shapefiles were not available.
    b. Google Earth Pro displays street segment lines with relative accuracy when viewing Street View panoramas, so determining the extent of, and the included trees for each

segment is much simpler and presumably more accurate (Figure 2.1). There are a number of free GIS programs which could be used to create shapefiles of this nature, albeit with a potentially large time investment depending on user ability with GIS software.

Figure 2.1: Top: view of municipal boundaries (white) and street segments (black) for the three surveyed communities, shown as shapefile lines and polygons imported into Google Earth Pro. Bottom Left: A street segment as seen from a GSV panorama. Bottom Right: Close up view of some street segments.

3. The virtual inventory was then conducted by estimating size and species for every municipal tree on each study segment by visual examination of GSV panoramas.
    a. Starting at one end of each study segment, I worked up one side of the street scrolling through the GSV panoramas, and then back down the other side, stopping at each tree

and recording estimations. This is similar to how one might inventory a street segment in the field.

b. Each tree was plotted as a point in ArcMap as close as possible to the location that it appears in GSV, by using simple visual clues or by locating the tree within the aerial imagery in ArcMap, if possible.

c. Tree data attributes (discussed below) were filled in on the file's attribute table in ArcMap. Domains (drop tables containing possible input options) were used for applicable attributes to expedite the data collection process. Other users may elect to use a different approach. For example, one could record tree locations in an Excel spreadsheet using the estimated street address provided on the Google Earth Pro interface by Google's address locator.

4. The trees in the original inventory were added to the ArcMap document as point features. These trees were then spatially joined to the virtually estimated trees to produce a new layer which contains trees that have attributes from both the field and virtual surveys.

    a. This joined layer was then evaluated on a tree by tree basis to ensure that all the joins made sense. Some trees which only occurred in one inventory were excluded from this matched dataset, but they were tracked to assess agreement in tree counts between the field and virtual surveys.

    b. The matched dataset was then exported to Microsoft Excel for analysis and reformatting for i-Tree Streets, following formatting instructions provided at https://www.itreetools.org/.

    c. All i-Tree project input options (area size, population, and utility prices) were copied from the original inventory. In a new virtual survey these values can be obtained by

measuring the area in Google Earth, using municipal information, and exploring census data.

d. Segments which did not contain any trees were entered into i-Tree manually so that the model would count these street segments in the total number of segments surveyed.
e. The output of this process is an i-Tree Streets project file which can display benefit reports for the tree survey data.

2.3. *Tree data collection*

To be specific about the visual estimation process, I am referring to one or more analysts physically looking through GSV panoramas and estimating values based on the appearance of the tree within the photograph. Species estimation was completed based on visual characteristics (Figure 2.2). For example, the leaf shape and bark characteristics for red oak trees is distinctly different than that of pears, and in addition pear leaves are usually rather glossy, and the growth form is often oval-shaped. However not all of these characteristics may be available in each panorama. Small trees often did not have any foliage close enough to the panorama center to discern distinct characteristics, whereas larger trees often had many branches growing out over the street space and towards the camera. And even when clear imagery is available, some distinctions may still be difficult. Different species within the same genus are often separated by only minute characteristics for which an arborist may normally rely on a hand lens, or interact with the tree directly to identify the species.

Figure 2.2: View of a tree from within a GSV panorama. This is a larger tree, so many of the leaves are close enough to the camera to be easily discernable. Based on the available view of the leaves, bark, and growth form, I would identify this tree as a silver maple (*Acer saccharinum*).

Diameter at breast height (DBH) was estimated using GSV by looking at the trunk or bole of the tree, at the standard measurement level of 4.5 feet off the ground. I found this to be more difficult when there was no good frame of reference as to the size of the tree, and easier when other objects of familiar size within the panorama could be used as comparison or a reference from which I could "place myself within the scene" mentally. Useful reference objects included anything within the scene that had a size which I was familiar with, some of which were people, cars, street signs, and telephone poles. Additionally, I found that comparing the base of the tree to the planting strip size was especially helpful, since most planting strips are of similar width from curb edge to sidewalk edge.

A number of attributes are required by the i-Tree model including a street segment identification number, DBH, and a tree identification to the species level. These attributes all have collection instructions defined in the i-Tree Streets User Manual accessible at http://www.itreetools.org. I collected all attributes which required numerical estimation in inches or feet, simply because of

familiarity with the unit sizes. In addition, I experimented with collecting several optional attributes which should have theoretically made the benefits assessment more accurate. Attributes like wood and leaf condition, which have a set number of options defined by the i-Tree model, were easy to collect and did not significantly increase the time spent on each tree; however, numerical attributes like tree height, height to crown base, and crown width (in two directions) dramatically increased per tree data collection time. These attributes were also significantly more difficult to estimate than DBH because they required me to scroll through the available panoramas on the street segment to find one with an adequate view of the height or width of the tree in question. The collection of these additional numerical estimations was dropped from the final study protocol because of the time required. If this methodology were practically applied for municipal management purposes, a few other variables could be collected including utility wire interference, sidewalk heave, maintenance needs, and other infrastructure conflicts that may need to be addressed, thereby increasing the utility and expanding the scope of the study from a benefits assessment only perspective. However, we did not test these variables so it is unclear how well they can be assessed in a GSV virtual survey. Lastly, I recorded some information at each tree about the available panoramic imagery including the date of the imagery, and its quality. This was helpful for comparing imagery to field survey dates and learning which panoramas provided the best identification results, or prevented accurate estimations.

2.4 *Calibration of Virtual Estimations*

After completing the first of three municipalities and comparing my estimations to the field measurements, I became dissatisfied with my DBH estimation accuracy. To increase the accuracy of GSV visual estimations I visited a subset of trees in the field. I picked several trees with varying species and size ranges from the local street tree population in Muncie, Indiana, made my DBH estimations using GSV, and then visited these trees, *in situ*, to verify accuracy and mentally calibrate my on-screen

estimation. In addition, I constructed a reference sheet containing GSV pictures of trees of known sizes, separated into common size classes (Appendix 2).

3. Proof of Concept

3.1 *Study Area*

This research utilizes both field and virtual survey data collected from the same geographic area for the purpose of comparison. The field survey, which was completed from August to October 2013, measured trees in nine communities in metropolitan Cincinnati, Ohio. These nine communities were selected to capture a range of geographic settings and socioeconomic conditions. The virtual survey, with imagery ranging from 2009 - 2016, revisited the field sites in three of these residential Cincinnati suburbs: Mt. Healthy (39.23° N, 84.55° W) Reading (39.23° N, 84.44° W), and Wyoming (39.23° N, 84.47° W). Mt. Healthy has a population of 6,601; Reading, 10,357; and Wyoming, 8,404. Of these three communities, only Wyoming has an active street tree planning program, but all three have street trees from past planting efforts (Berland and Lange 2017).

3.2 *Methodology*

In each community 10% of all local city street segments were selected at random, from all available street segments using U.S. Census Bureau's 2013 TIGER/line roads data. The 10% sample size is greater than the 6% sample suggested by i-Tree Streets, which will help reduce the standard error of the subsequent benefits estimation. The random selection process was initially completed for the field survey, and then the same street segments were reused for the virtual survey to form the comparative basis. In the field, each segment was visited and street trees which appeared to be a part of the public easement were inventoried according to i-Tree Streets protocols. In the three communities selected for the virtual survey, the same street segments were revisited and tree attributes were visually estimated.

I-Tree benefits estimation model works by comparing collected tree data to growth curves from trees measured in regional reference cities across the United States. Tree species codes are used to match trees to the appropriate growth curve, and then diameter is used as a means to estimate leaf area. These curves were produced by measuring trees spanning a range of diameter values and estimating their crown size and leaf area by processing computer images of the tree (i-Tree 2010). This means that the model is mainly driven by available species growth curves and diameter values for leaf area estimation. I-Tree Streets requests a number of values for benefits assessment including a species identification, DBH, and wood and leaf condition. Of these metrics, only tree species, DBH, a segment number, and some basic information about the study area are required for the model to calculate benefits, it is therefore very important to accurately estimate species and diameter values within the virtual survey to provide comparable results.

The virtual survey data collection was completed by a single analyst who held a Bachelor of Science in field botany and had prior experience identifying trees and collecting basic forestry measurements. After completion, the collected tree data was matched to field observations. This was done by spatially linking the two datasets using ArcMap, and then manually checking the matched dataset for discrepancies. Temporal differences between GSV panorama capture dates and the field survey caused issues at locations where trees were planted, or underwent maintenance, removal or replacement between the time when the field and virtual surveys were completed. As a result, the matched dataset contains only those trees that perceivably existed at the time of both field and virtual survey completion. Species and size class agreement within the matched dataset were evaluated using Cohen's kappa, a statistic that determines whether the level of agreement between virtual and field survey values is greater than if the virtual values were produced by random chance from available choices. Diameter agreement was assessed using a linear regression model, where a line slope of 1

would indicate perfect virtual estimation accuracy, and where the coefficient of determination (R^2) was used to evaluate the consistency of virtual estimations.

The matched data was then exported to Microsoft Excel to be reformatted for benefits analysis in i-Tree Streets. The reformatting process includes assigning the i-Tree species codes to the species recorded in each study. Cincinnati is part of the lower Midwest region, so the species codes from the lower Midwest were used. Not all species have a corresponding species code, so some substitutions had to be made (e.g., there was no code for *Quercus acutissima*, so the more generic i-Tree code for *Quercus* spp. was used instead). In addition, street segments which contained no trees were added to the spreadsheet using a placeholder species code which indicates no tree presence. This is to ensure the model accounts for all street segments and not just those which contained street trees.

The i-Tree Streets model contains two methods of estimating tree benefits, sample and complete inventories. A sample inventory is intended for surveys which collected data from only a portion of the total street segments for a given area, and upon benefits analysis will automatically extrapolate results to the entire study area, with the extrapolation based on the total number of street segments versus those sampled. A complete inventory assumes that all streets within the study area were surveyed, and calculates benefits for only those trees entered into the model, with no extrapolation. For reasons previously discussed, I intentionally omitted a portion of both the field and virtual survey's trees from the matched datasets; for this reason, benefits analysis of the field and virtual matched datasets was done using the complete inventory setting in i-Tree Streets, so that only the 558 matched trees are included in the analysis, and citywide extrapolation of results does not introduce additional error due to the intentional omission of records. I also included sample inventory analysis of the full field and virtual datasets, without any records omitted, so that the virtual survey, as a methodology, may be compared to the field survey. For the remainder of this document, when I refer to

sample and complete inventories, I will be referencing the i-Tree Streets analysis settings, and not common survey practices.

3.3 *Survey Results*

I made visual estimations of a total of 606 street trees along 115 study segments totaling 15.8 km of street length. Completion of two municipalities, Reading and Wyoming, took around 11.2 hours; time was not recorded for Mt. Healthy. The field survey captured a total of 597 street trees along the same street segments. Between these two surveys, 558 trees (93.4% of the total dataset) were positively matched to create a matched dataset in which trees contain both a field and virtual record. Of the matched tree records, the level of agreement between the field and virtual surveys was 89% for genera (kappa = 0.86; $p < 0.001$) and 58% for species (kappa = 0.56, $p < 0.001$). Genus agreement was better for trees with larger diameters, and for distinct genera with few specimen (Figure 2.3, Table 2.1). Species agreement was higher for genera which contained fewer species options within the study area (Table 2.2). Linear regression shows relatively consistent underestimation of diameter throughout the virtual survey (Figure 2.4). Trees in the field survey had larger recorded diameters 52.5% of the time. When separating DBH values into size classes (1-3” 3-6” 6-12” 12-18” 18-24” 24-30” 30+), 67% of all matched trees are correctly placed in the same category (weighted kappa = 0.73, $p < 0.001$). In addition, diameter estimation accuracy improved as study areas were completed. Mt. Healthy, the first of the three municipalities had the largest diameter underestimation issue. After the completion of Mt. Healthy I implemented a size reference guide (Appendix 2), and this, along with the accumulated experience may account for the increased accuracy in Reading and then further in Wyoming (Figure 2.5).

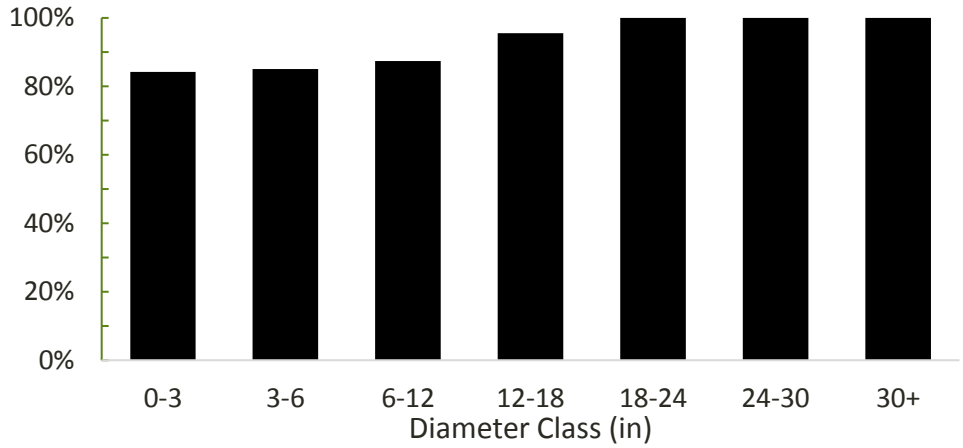

Figure 2.3: Genus agreement by diameter class.

Genus	Field Count	Virtual Count	Matches	Agreement
Zelkova	27	27	27	100
Cercis	6	6	6	100
Liquidambar	5	5	5	100
Celtis	3	3	3	100
Liriodendron	3	3	3	100
Aesculus	2	2	2	100
Catalpa	2	2	2	100
Juglans	1	1	1	100
Picea	1	1	1	100
Pinus	1	1	1	100
Platanus	1	1	1	100
Quercus	41	40	40	97.6
Acer	147	149	145	96
Gleditsia	35	37	35	94.6
Pyrus	80	88	79	88.8
Tilia	8	7	7	87.5
Malus	25	22	20	74.1
Fraxinus	85	62	62	72.9
Prunus	21	20	17	70.8
Syringa	14	18	12	60
Amelanchier	5	7	4	50
Ulmus	41	48	29	48.3

Table 2.1: Genus agreement between field and virtual surveys, values of 0 not shown. Agreement calculated as: Matches/(Field Count + Virtual Count – Matches)

Species	Field Count	Virtual Count	Matches	Agreement
Zelkova serrata	27	27	27	100
Cercis canadensis	6	6	6	100
Liquidambar styraciflua	5	5	5	100
Celtis occidentalis	3	3	3	100
Liriodendron tulipifera	3	3	3	100
Acer palmatum	2	2	2	100
Catalpa speciosa	2	2	2	100
Juglans nigra	1	1	1	100
Quercus acutissima	1	1	1	100
Gleditsia triacanthos	35	37	35	94.6
Pyrus calleryana	79	88	78	87.6
Acer saccharinum	33	31	29	82.9
Malus spp.	25	20	20	80
Acer rubrum	43	49	37	67.3
Quercus rubra	16	12	11	64.7
Syringa reticulata	14	18	12	60
Acer saccharum	36	53	33	58.9
Tilia cordata	4	7	4	57.1
Acer platanoides	16	7	7	43.8
Acer campestre	15	3	3	20
Quercus alba	2	16	2	12.5
Quercus bicolor	12	1	1	8.3

Table 2.2: Species agreement between field and virtual surveys, agreement values of 0 not shown. Agreement calculated as: Matches/(Field Count + Virtual Count – Matches)

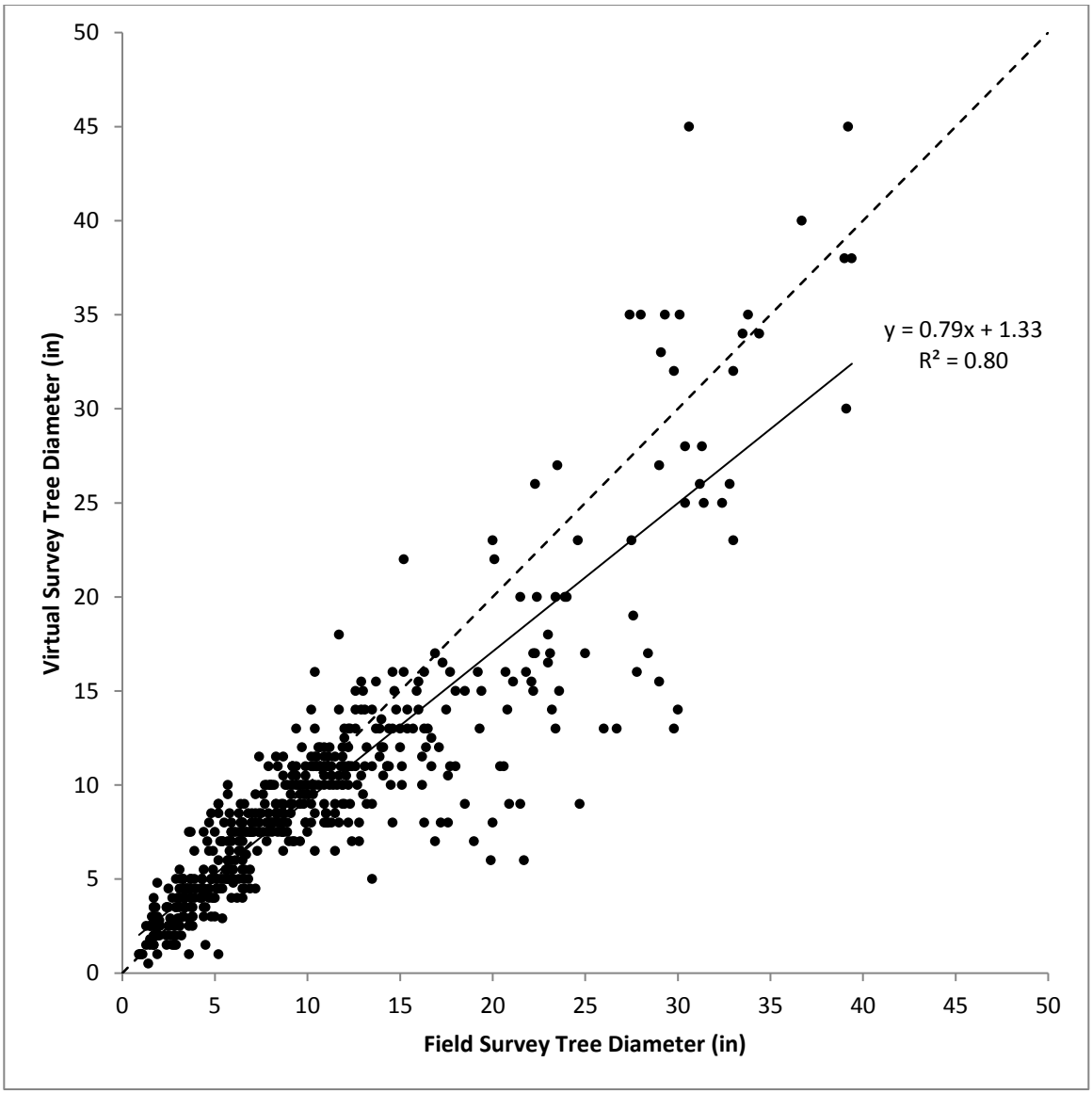

Figure 2.4: Diameter comparison between field and virtual tree diameters, where virtual estimations are plotted against field measurements with the dashed line representing 100% accuracy. This graph includes all three study municipalities.

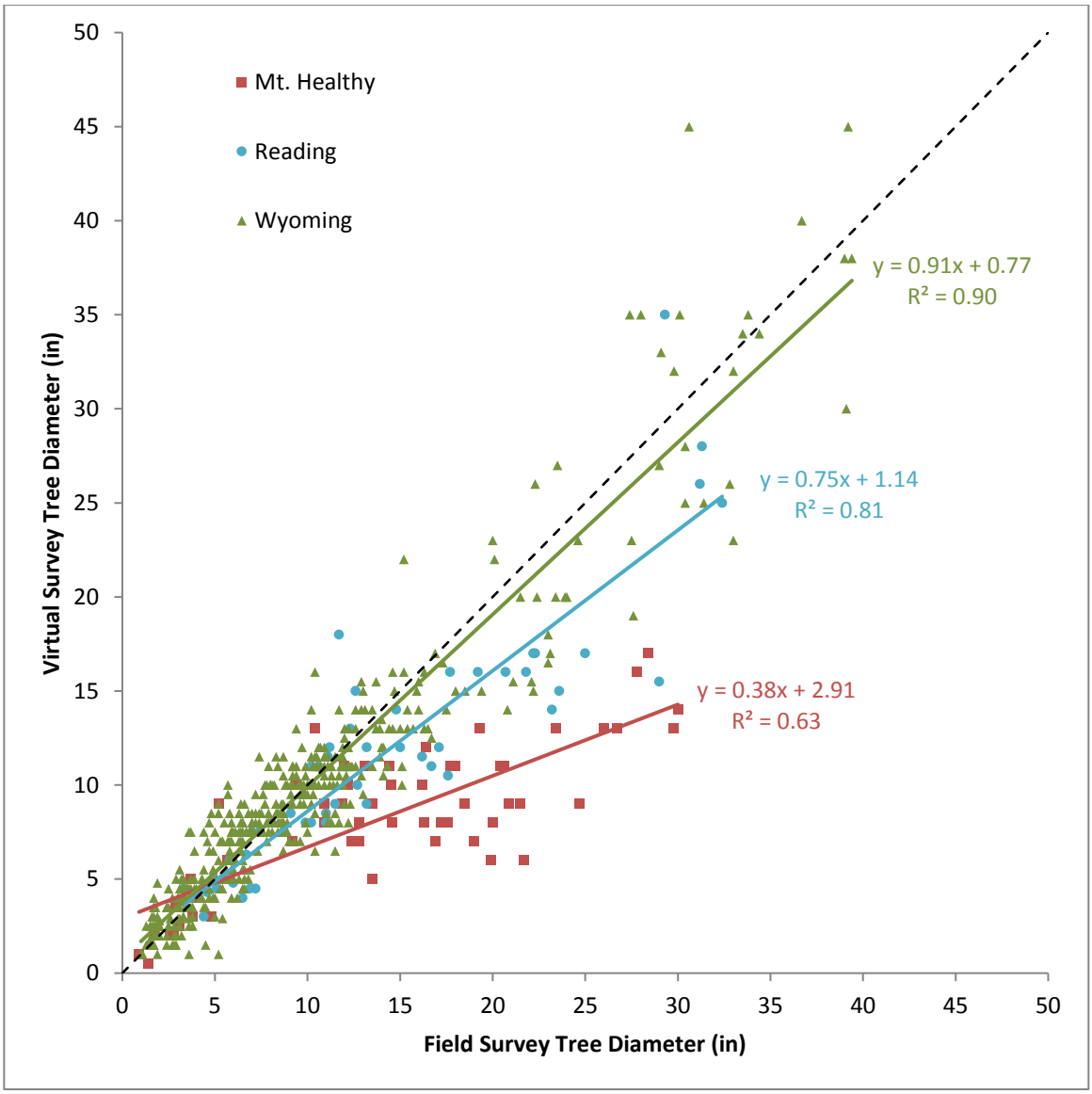

Figure 2.5: Diameter comparison between field and virtual tree diameters, where virtual estimations are plotted against field measurements with the dashed line representing 100% accuracy. The study municipalities in this graph are separated to show how diameter estimations improved over time from Mt. Healthy to Reading to Wyoming.

3.4 *I-Tree Benefits Assessment and Results*

After formatting the matched dataset for i-Tree Streets import, 62.7% of the species input codes were the same between the field and virtual surveys. Complete inventory benefits analysis of the matched dataset shows total net annual benefits were higher for the virtual survey by $7, with energy, air quality, and stormwater values greater than field survey estimates, and CO_2 and aesthetics lower (Table 2.3). The order of major contributing species differed between field and virtual surveys, but many of the key species remained the same (Table 2.4). Sample inventory analysis of the full datasets shows 5,295 (±832) trees extrapolated for the field survey and 5,375 (±833) for the virtual survey, with net annual benefits higher for the field survey by $80,752 (29.3% difference) (Table 2.3).

	Field		Virtual	
Benefits	Sample	Complete	Sample	Complete
Energy	46,124 (±7,740)	2,618	21,761 (±3,581)	2,728
CO_2	9,144 (±1,549)	410	6,040 (±1,025)	386
Air Quality	5,639 (±951)	334	2,663 (±436)	349
Stormwater	147,814 (±25,035)	8,050	61,774 (±10,387)	8,503
Aesthetic/Other	106,927 (±19,328)	4,913	142,660 (±24,588)	4,366
Total	315,649 (±53,767)	16,325	234,897 (±39,377)	16,332

Table 2.3: Net annual benefits for field and virtual surveys in U.S. dollars. Standard error shown where applicable. Sample values represent the sample inventory of the full dataset; compete values represent the complete inventory of the matched dataset

Energy

Field		Virtual	
Oak	18.3	Oak	20.8
Sweetgum	16.3	Green ash	17.2
Green ash	15.6	Sweetgum	17.2
Elm	14.9	Tulip tree	17.2
Boxelder	14.0	Elm	15.6
Japanese zelkova	13.2	Japanese zelkova	14.8
Chinese elm	13.1	Ash	14.4
White ash	12.3	Norway maple	13.5
Tulip tree	11.3	Scarlet oak	12.4
Maple	11.2	Hedge maple	12.1

CO_2

Field		Virtual	
Oak	6.8	Oak	6.3
Tulip tree	3.0	Scarlet oak	5.1
Elm	2.7	Elm	2.6
Chinese elm	2.5	White oak	2.6
Sweetgum	2.3	Northern hackberry	2.6
Northern red oak	2.2	Northern red oak	2.3
Green ash	2.2	Silver maple	2.2
Silver maple	2.1	Hedge maple	1.9
Maple	2.0	Maple	1.7
Hedge maple	1.7	Norway maple	1.5

Air Quality

Field		Virtual	
Sweetgum	2.1	Sweetgum	2.2
Elm	2.1	Tulip tree	2.2
Green ash	2.0	Green ash	2.2
Chinese elm	1.8	Elm	2.2
Oak	1.8	Oak	2.0
Japanese zelkova	1.7	Japanese zelkova	1.9
Boxelder	1.7	Ash	1.9
White ash	1.6	Norway maple	1.7
Tulip tree	1.4	Hedge maple	1.4
Norway maple	1.4	Northern catalpa	1.3

Stormwater

Field		Virtual	
Oak	64.5	Oak	79.6
Elm	54.4	Elm	57.2
Chinese elm	47.2	Sweetgum	48.5
Boxelder	47.2	Tulip tree	48.5
Sweetgum	45.8	Green ash	48.5
Green ash	43.6	Norway maple	44.0
Eastern white pine	40.9	Japanese zelkova	41.4
White ash	39.3	Ash	40.4
Sugar maple	37.4	Hedge maple	39.1
Japanese zelkova	36.4	Sugar maple	37.9

Aesthetic

Field		Virtual	
Siberian elm	57.9	Northern hackberry	48.9
Northern hackberry	53.4	Scarlet oak	48.7
Oak	50.5	Oak	46.4
American elm	47.5	White oak	39.8
Northern red oak	38.3	Northern red oak	38.8
Chinese elm	35.8	Maple	36.3
Swamp white oak	34.5	Black locust	34.0
Elm	33.2	Elm	33.2
Tulip tree	32.6	Silver maple	27.2
Chinkapin oak	31.5	Japanese maple	26.8

Total

Field		Virtual	
Oak	141.9	Oak	155.1
Elm	107.2	Elm	110.8
Chinese elm	100.4	Scarlet oak	103.9
Maple	81.0	Northern hackberry	91.7
American elm	78.7	Hedge maple	79.0
Tulip tree	78.4	Maple	74.1
Sweetgum	75.0	Norway maple	72.6
Hedge maple	74.8	Sweetgum	68.9
Green ash	74.0	Tulip tree	68.9
Northern hackberry	72.5	Green ash	68.8

Table 2.4: Top ten species in each i-Tree benefits category for matched field and virtual complete inventory benefits analysis. Values represent net annual benefits by species in U.S. dollars.

3.5 *Discussion and Conclusions*

This study explored the use of GSV as a source of freely available ground level imagery for characterizing street trees and the environmental benefits they provide. Results of the virtual survey show potential for the use of GSV to conduct virtual street tree surveys. As this was a pilot study, a number of decisions were made that likely had alternative solutions. I chose to carry out many of the required processes in ways that I was capable of and made sense to me, but may have not been the most efficient or effective for practical applications. For example, I entered data into ArcGIS tables for the spatial placement of tree points; however, data entry could also be done through i-Tree Streets directly, or through any spreadsheet program. In addition, I collected those variables which suited the needs of my comparative study, and mainly relied on DBH and species identification as model parameters. As mentioned previously, the variables collected could be tailored to fit the survey needs.

Using GSV imagery as a source for ground based remote sensing data has a number of potential applications, as discussed in the previous chapter; however, there are also a number of issues with the GSV panoramas. Some panoramas include image stitching distortions. I mostly saw this near the top or bottom of the panorama, so this was not much of an issue for DBH estimation, but it sometimes affected my view of foliage. From what I have read, this could be remedied by requesting the unstitched image portion from the GSV API. Additionally, censored sections occasionally impacted the view of the study tree. On one street segment, a homeowner had apparently requested his/her entire home be censored, largely blocking my view of the street trees for that property parcel (Figure 2.6). In the United States, Google's censoring algorithms are supposed to obscure faces and license plates, but occasionally blur other objects mistakenly (Frome et al. 2009). In other countries users may see more censoring or GSV coverage issues due to government restrictions (Rakower 2011). Another concern is objects physically between the panorama camera and the trees (Figure 2.6). I experienced this several times in the form of vehicles blocking the view of the tree, and I suspect this problem would become a greater

nuisance in denser urban areas, or places where street side parking is common. Lastly, I experienced an issue with imagery dates while conducting my survey. Panoramas from different parts of the study area often had different imagery dates, which were sometimes before and sometimes after when the field survey was conducted (Figure 2.7). This has implications for any virtual survey conducted using GSV imagery because the survey will contain missing values for a specific date range, which makes analysis of change over time very difficult. Recording the dates of the panoramas helped me work through most of my time-related issues when matching trees between the field and virtual surveys, and while any practical applications will not be referencing a field survey, I do recommend keeping track of the imagery dates because of this temporal nuance to the imagery format.

Figure 2.6: Trees can be obstructed from view physically by a vehicle or blurred by Google's censoring software. Obstructions may be a greater concern in urban core areas.

Figure 2.7: Changes occur to a city's tree population between imagery dates. In the images, above, the left image is from the field survey in 2013 and shows a tree in the process of being removed, and in the right image from Street View in 2014 the tree is gone. This occurred several times throughout the study, where a tree which may have even been healthy in the field study was simply missing from GSV panoramas.

While I originally thought that many of the differences between the virtual and field surveys were driven by the underestimation of DBH, after running the i-Tree analysis on the matched dataset, I believe that accurate species identification is just as, if not more important. The virtual and field surveys had 63% of the same i-Tree species codes, and the field survey trees had larger diameter values 53% of the time. Even so, the complete inventory benefits analysis for the matched virtual survey had higher total net annual benefits, by a small margin. I-Tree benefit estimations are driven largely by DBH values, but use growth models from several available species to relate DBH to leaf area for the calculations (i-Tree 2010). Some the virtually surveyed trees which were misidentified may have used growth models that have more leaf area in the analysis. One major contributor to this difference may have been the ash

trees; in the virtual survey I identified almost all ash trees as *Fraxinus* spp. as opposed to guessing between *F. pennsylvanica* and *F. americana*. However, the field survey differentiated between these species (Table 2.2). In addition, it looks as though the DBH overestimation occurred mainly for smaller size trees, which may have had a larger impact on total leaf area than the medium to large size trees which I mostly underestimated (Figure 2.4). Although it is hard to pinpoint the exact set of values that led the matched virtual survey to have higher benefits, I believe that the growth models of mismatched species are a large contributor. As for the sample inventory analysis of the full datasets, the field survey values are much higher. The virtual sample did extrapolate about 80 more total trees, but the DBH underestimation likely applies more in this sample inventory scenario, and accounts for the large difference in modeled benefits.

One main concern for conducting the virtual survey was accurate species identification through visual interpretation of GSV panoramas. In this study, I found agreement between field and virtual survey identification was more accurate for larger trees. Trees of smaller diameter were often further away in the imagery (in addition to being physically smaller) and thus were represented by fewer pixels, leading to more ambiguity about distinguishing characteristics, some of which may or may not have recognizably developed yet because of the tree's young age. Species identification was also influenced by the number of species within a genus, because it is more difficult to distinguish between similar species within the same genus. This led to a higher species agreement within single species genera like *Gleditsia, Pyrus, and Zelkova*, and a much lower species agreement in those genera with many common species like *Acer, Quercus, and Ulmus* (Table 2.2). In relation to the i-Tree benefits estimation, genera may be an acceptable level of identification, as many of the growth forms for specific species are not represented in the i-Tree model. Furthermore, if benefits analysis was the only objective, species identification could potentially be largely eliminated by matching trees to available growth forms within

the model, though this would likely still require some amount of dendrological knowledge and some detailed knowledge of the growth models.

Diameter estimation was my second main concern for the completion of the virtual survey. In this study, I visually estimated the diameter of each tree with what amounted to an educated guess. There are several context clues within GSV images that can hint towards diameter, and I have had experience physically measuring trees, but there was no standard methodology to the diameter estimation. I found the reference sheet to be helpful, and experience led to better estimations. I would however expect different estimation accuracy from a different rater. I had more underestimations than overestimations, but another person could easily overestimate more on average leading to inflated environmental benefits. Averaging multiple raters is one way to potentially improve diameter estimation accuracy, but would require more manpower and time. Diameter estimation accuracy could also be potentially increased with the use of photogrammetry software, and is a topic for future research inquiries. In this study, I estimated diameter in inches, but i-Tree only requires that diameter is entered in size class ranges. This would reduce the number of input options for tree sizes, and may lead to higher accuracy at the expense of precision.

Further concern is raised when considering inter rater differences. I came into this study with experience measuring and identifying trees, but ideally the study could be performed by anyone. Creating a reference sheet containing imagery of common species and their distinguishing characteristics or pictures of tree boles of trees of known size within the GSV panorama setting may improve identification and diameter estimation success among less experienced virtual survey analysts. Additionally, evaluating a test set of trees of known sizes for calibration purposes prior to conducting the full survey may help to normalize rater estimations. I found these quality control methods helpful for improving my GSV tree estimations, and I believe this is reflected in my results (Figure 2.5). In any practical application of this survey technique I strongly recommend completing this sort of field-

checking calibration as well as creating some sort of GSV panorama based reference or completing some other form of quality control before beginning the virtual survey.

This method of generating estimates of tree benefits may be useful to entities which cannot afford the time and monetary investment that a traditional survey requires; however, the relatively coarse tree data generated by this survey technique does not replace the need for arborists conducting detailed assessments of tree health and maintenance needs in the field. An understanding of the municipal forest structure can help urban forest planners make educated management decisions. The virtual survey was completed in a much shorter time and at reduced cost compared to the field survey, and used freely available imagery and benefits modeling software. The i-Tree benefits were closely related when considering only results from matching trees, but large differences were seen when analyzing the full datasets. Tree species identification was markedly more difficult through the GSV panoramas than traditional *in situ* evaluation, but provided a baseline understanding of the forest structure. The reduced species accuracy may be an acceptable level of error given the reduced study time. The tree diameter estimation accuracy is likely to change largely from rater to rater, and is a topic for future research inquiry. In this study, diameter estimation improved as I gained experience with the technique. A reference sheet and a calibration dataset may have aided the increased accuracy and are tools to consider in similar future studies. One key downside to the use of GSV imagery is that the availability of imagery is limited to dates provided by Google, and imagery dates are likely to vary throughout the study area. Overall, it might be advantageous to use virtual study data when considering its shortcomings in data accuracy, depending on the intended use of the data set.

CH. 3. EXPLORATIONS IN WETLANDS DETECTION USING FREELY AVAILABLE DATA PRODUCTS

Abstract

Wetlands are important natural resources that often contain unique topographies, soils, and water features, as well as a high number of threatened and endangered plant and animal species. Tracking wetlands can involve extensive fieldwork and coordination with private landowners to locate and visit possible wetland sites for determination. The Robert Cooper Audubon Society (RCAS) identified the need to locate ecologically valuable wetlands within their service area in East Central Indiana. Their existing map product uses outdated input data, displays with a coarse resolution, and is not specific to wetlands detection. Updated and more recent data products including 4-band National Agricultural Imagery Program (NAIP) aerial photography, light detection and ranging (LiDAR) point clouds, National Wetlands Inventory (NWI) polygons, and soil types can be used alongside remote sensing techniques to identify possible wetland areas on a finer scale. This project created a new map product for RCAS that aggregated several determining factors and data layers into remote wetlands detection efforts, with the goal of creating a map product with greater accuracy and precision. Moving forward, validity of outputs should be verified in the field using wetlands determination methods. This map product can potentially be used to focus wetlands conservation efforts with a more meaningful impact.

1. Introduction

1.1 *Overview of Wetlands*

The term "wetlands" has many definitions because wetlands encompass a great variety of landforms, environments, organisms, and hydrologic systems. In general, a wetland is an area in which the hydrology is the main determinant of the type of ecosystem that develops there (Cowardin et al. 1979). The U.S. Fish and Wildlife Service's official definition of wetlands from Cowardin et al. (1979) is as follows:

> *"Wetlands are lands transitional between terrestrial and aquatic systems where the water table is usually at or near the surface or the land is covered by shallow water. For purposes of this classification wetlands must have one or more of the following three attributes: (1) at least periodically, the land supports predominantly hydrophytes; (2) the substrate is predominantly undrained hydric soil; and (3) the substrate is nonsoil and is saturated with water or covered by shallow water at some time during the growing season of each year."*

Wetlands are valuable natural resources that provide many ecological services such as floodwater storage, improved water quality to surrounding areas, climate regulation on a regional scale; and, most notably, they support a rich biodiversity that is not found elsewhere which provides habitat and spawning grounds for a plethora of threatened and endangered species (Gallant, 2015; Ozesmi and Bauer, 2002; U.S. Fish and Wildlife Service, 2002) . Since the United States was first settled by Europeans, there has been a nationwide decline in wetlands of over 50%, with only around 110.1 million acres left in 2009 (Dahl 2000; 2011). There are many factors that have led to the steady decline in wetlands over the last few centuries. Some wetlands are utilized for economic activities, such as reed collecting and peat moss harvesting; but, urban and rural development and land conversion for farming purposes continue to be the greatest threats to wetland preservation. Additionally, climate change is likely to disturb many delicate wetland ecosystems (Johnson et al. 2005). However, more and more wetlands are becoming protected through laws, regulations, and conservation measures. Protected lands are those that are managed for the continued conservation of their ecosystem. Managing

protected lands traditionally consists of considerable field work to identify areal extent and determine appropriate conservation strategies; however, field work is often time-consuming, costly, and may not allow for the coverage of the entire interest area. Remote sensing techniques for the management and conservation of protected lands are becoming increasingly sought after and useful tools (Wang 2011).

1.2 *Wetland Remote Sensing*

Remote sensing of wetlands is advantageous as wetlands are often sensitive to human intrusion and relatively inaccessible. Satellite remote sensing of wetlands allows for greater coverage, analysis of change over time, and is less costly and time consuming than aerial photography. On the other hand, aerial photography often has a higher resolution, and is not limited by cloud cover or satellite return time. One limiting factor in wetlands detection is the huge range of diversity of wetland types. Wetlands can be permanently flooded, seasonally flooded, or contain wet soils but have no above-ground water. Vegetation types vary by wetland morphology and spectral signatures (i.e., the reflectance of energy at various wavelengths that is characteristic of a particular type of vegetation) change throughout the growing season. In general, it seems that wetlands containing more water for longer are easier to delineate using remote sensing techniques than those that only contain water for a small portion of the year, yet almost all wetland types have been studied to some degree using satellite remote sensing (Ozesmi and Bauer 2002).

Many different techniques have been used to identify wetlands using remotely sensed data. Visual interpretation is probably the most basic method; it is most suited to aerial photography and is very time consuming. One study used aerial image interpretation to detect the impact of human activity in the Balçovas' delta by examining land use over time (Bolca et al. 2007). The authors found that wetlands and natural areas in the delta had decreased by nearly 2000 ha from 1957-2005 due to rapid urbanization. Other methods of wetlands image analysis mainly rely on computational algorithms.

Unsupervised analysis methods have been popular for wetlands identification. In this approach, an image clustering algorithm groups pixels together based on their spectral values and then those groups are identified by the analyst (Ozesmi and Bauer 2002). Principal component analysis is a preprocessing step which can be completed prior to unsupervised classification to reduce the number of spectral bands before clustering and is therefore useful on datasets with many bands to work with. These methods have been used extensively and are relatively accurate when applied appropriately. Supervised classification methods rely on training data acquired beforehand to calibrate the analysis before applying the clustering algorithm to the area of interest. Hybrid classification methods use a supervised and unsupervised classifications in conjunction with the goal of improving overall accuracy, which can be useful for wetlands identification because of the great variability in wetlands types (Ozesmi and Bauer 2002).

In addition, several vegetation indices also exist that can aid in wetlands differentiation. The normalized difference vegetation index (NDVI) makes vegetation more distinguishable from non-vegetation using infrared and red imagery bands (Rouse et al. 1973). NDVI ranges from -1 to 1, where lower values represent open water, values near 0 indicate bare earth, and higher values indicate healthy, abundant vegetation. Soil adjusted vegetation index and optimized soil adjusted vegetation index function similarly to NDVI but attempt to eliminate variation in vegetation index values due to underlying soil types and their differences (Huete 1988; Rondeaux et al. 1996). There are also simple ratio and differential vegetation index, both of which also utilize differences in vegetation reflectance between the red and infrared bands (Pearson and Miller 1972; Richardson and Everitt 1992). Davranche et al. (2010) utilized a number of these indices alongside moisture indices to separate reed grasses from submerged vegetation in southern France using a binary classification tree algorithm, finding the algorithym to be 98% successful at finding areas of reed grasses, and 97% successful in locating submerged vegetation. Ancillary data has been used by many researchers to improve the accuracy of

wetlands classification. Data like soil, elevation, and tree cover have all been used as supplemental information to improve satellite remote sensing techniques for wetland identification (Ozesmi and Bauer 2002). One technique, employed by Wu et al. (2015), involves using topographic contour lines generated from elevation files and an algorithm built to mimic human map interpretation to identify and map the extent of surface depressions. This methodology was utilized in Wu and Lane (2016) to identify surface depressions in the Prairie Pothole Region of North America. The authors were able to identify over 12,000 topographic depressions structures in the 506 km^2 study area, producing depression locations with greater accuracy and precision than the existing NWI maps for the area. Other studies have employed various methodologies. Levick and Rogers (2006) used an object based image analysis to delineate woody vegetation in Kruger Park, South Africa, and used height and crown size extracted from LiDAR to further differentiate between vegetation types. To provide another example, Dvorett et al. (2016) used Landsat images from 1994 to 2011 to find inundated wetlands in the Pleistocene Sand Dunes Ecoregion of Oklahoma, identifying water pixels using a decision tree analysis method. The authors were able to identify about 700 more ephemeral wetlands than NWI maps because of their multi-year approach.

1.3 *RCAS Wetlands Mapping*

The Robert Cooper Audubon Society (RCAS) began as an independent society for Delaware County, IN in the 1960s, later becoming a regional chapter of the National Audubon society and extending its operating range to Blackford, Grant, Henry, Jay, Madison, and Randolph Counties (RCAS n.d.). RCAS is a nature focused group, with an emphasis on conservation and education. One conservation interest is the preservation of ecologically valuable wetlands sites. Often these locations can be small and undocumented by available resources, thereby requiring considerable field work to identify locations and develop conservation strategies. The current map used for reference during RCAS conservation

efforts is not specifically tailored to wetlands detection, and instead shows the conservation potential of land parcels (see map at http://delaware.iaswcd.org/). This exploratory survey serves to generate high resolution mapping products from freely available data which can be used as an aid to identify potentially valuable wetlands sites. This aid can then be used to better focus conservation efforts by identifying new wetlands sites to explore and protect.

2. Methodologies

2.1 *Study Area*

RCAS operates in Blackford, Delaware, Grant, Henry, Jay, Madison, and Randolph Counties, IN. For the purposes of this exploratory study I used data from Delaware and Henry Counties. As a Muncie, IN resident of around six years I have some familiarity with Delaware County. I have taken a number of botany courses in which I have visited wetland and natural areas around the county, and in working for Ball State University's Landscape Services as a GIS analyst I have become familiar with the campus grounds. In addition, RCAS has provided a hand drawn map of Henry County containing known wetland locations. My familiarity with Ball State's campus and the hand drawn map helped me spot check my results as I explored data analysis options and ensured data products aligned with intended outcomes.

2.2 *NDVI Vegetation Extraction combined with LiDAR Elevation*

2.2.1 *LiDAR Processing*

In the first portion of the project, I used LiDAR data to model the elevation of the bare earth surface and the heights of objects on the earth's surface. LiDAR data consists of millions of laser imaging points collected from airplane flights over the area of interest. These laser points contain elevation and signal strength, which can be used to generate high resolution elevation and surface models. Indiana's statewide LiDAR data was completed over a 3-year period from 2011-2013 with the LiDAR for Henry

County flown in 2012, and LiDAR for Delaware County provided from earlier data collected in 2008 (ISDP 2017). LiDAR data was downloaded for Delaware and Henry Counties from Indiana Spatial Data Portal (ISDP) in .las format. Processing LiDAR data for combination with extracted vegetation involved three major steps: (1) creating a digital elevation model (DEM), (2) creating a digital surface model (DSM), and (3) combining them into an elevation raster file. DEM files show the elevation of the bare earth, and are created using only those LiDAR points which return from the ground. The DEM excludes buildings and vegetation and can be utilized for identifying depressions which may contain a wetland. The DSM is generated similarly but includes all point returns and therefore contains buildings, trees, and other objects on the landscape. Subtracting the DEM from the DSM can return the distance from the top of objects to the bare earth. This produces an image file containing heights (in feet) of objects in the county, including buildings, trees, and low-lying vegetation. At the county level, the LiDAR data contains a massive number of total points (940 million for Delaware County, 1.5 billion for Henry County). I was unable to create elevation rasters at the county scale directly due to software limitations and the large number of LiDAR points; instead, I processed each sub-county LiDAR file individually, iterating through each file in the county (Figure 3.1). After elevation rasters were created for each LiDAR file, I created a mosaic and added in all of the elevation rasters; this product can then be joined back to the extracted vegetation to create the final elevation layers (Figure 3.3)

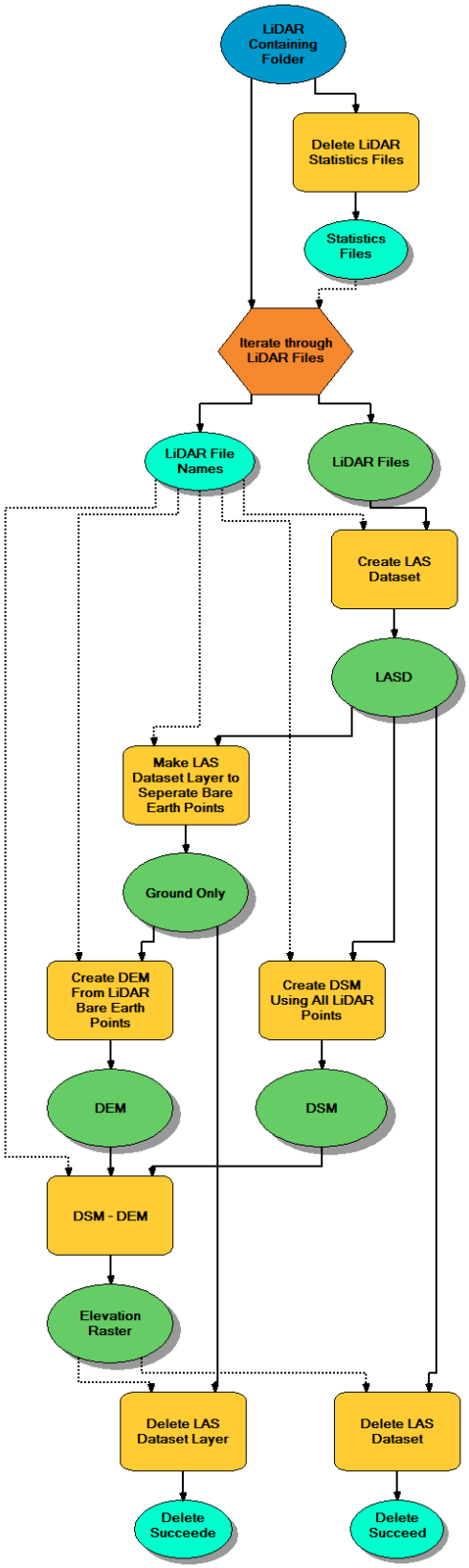

Figure 3.1: Processing steps for creation of DEM, DSM and elevation raster outputs from individual LiDAR files

2.2.2 *NAIP Imagery Processing and Final Output*

NAIP imagery has been widely used for a variety of purposes in government, private sectors, and education, and has been the basis for a number of publications (Aerial Photography Field Office 2008). NAIP is produced by the United States Department of Agriculture's Farm Service Agency, and has been providing leaf-on aerial photography since 2003 (USDA's FSA 2017). NAIP Imagery began on a 5-year acquisition cycle, but switched to a 3-year cycle in 2009. NAIP files are distributed as georeferenced tagged image file format (GeoTIFF) raster files with one meter resolution which represent one quarter of a USGS topographic quadrangle. NAIP imagery is collected by airplane, and is flown in the summertime to capture information about the agricultural growing season. Since 2007 the NAIP Imagery products for some states have been collected in four bands (the traditional red/green/blue bands plus a near infrared (NIR) band) (USDA's FSA 2017). For this project, I downloaded all of the 2014 NAIP quarter quad TIFF files for Delaware and Henry Counties, IN from ISDP (http://gis.iu.edu/). NAIP image tile sets from each county were merged into county wide GeoTIFF files.

NDVI was calculated using the merged county NAIP with the Raster Calculator tool in ArcGIS 10.3 (Equation 1). The result of this calculation was a raster layer with values ranging from -1 to 1, with values closer to 1 more likely to be vegetation. All values below 0 were then removed from this raster to exclude non-vegetation from further analysis (Figure 3.3).

("%Infrared_Band%" - "%Red_Band%") / ("%Infrared_Band%" + "%Red_Band%" + 0.00)

Equation 3.1: Raster Calculator formula for NDVI

The last step was to combine the extracted vegetation with the elevation raster to create a product which displays the heights of all the vegetative surfaces (Figure 3.3). This output was run through a low pass filter, which applies a smoothing effect using a 3 x 3 pixel averaging grid. Although this filter introduced error by artificially lowering vegetation heights, this was done to improve visual data readability, as the unfiltered tree canopies have a stippled appearance due to uneven LiDAR point

interception by foliage. This product can then be used to separate vegetative surfaces based on height (e.g. trees can be separated from grass and shrubs). The vegetation with elevations output was evaluated for accuracy by visual interpretation of a random sample of 100 points using accuracy assessment techniques outlined by Congalton (1991).

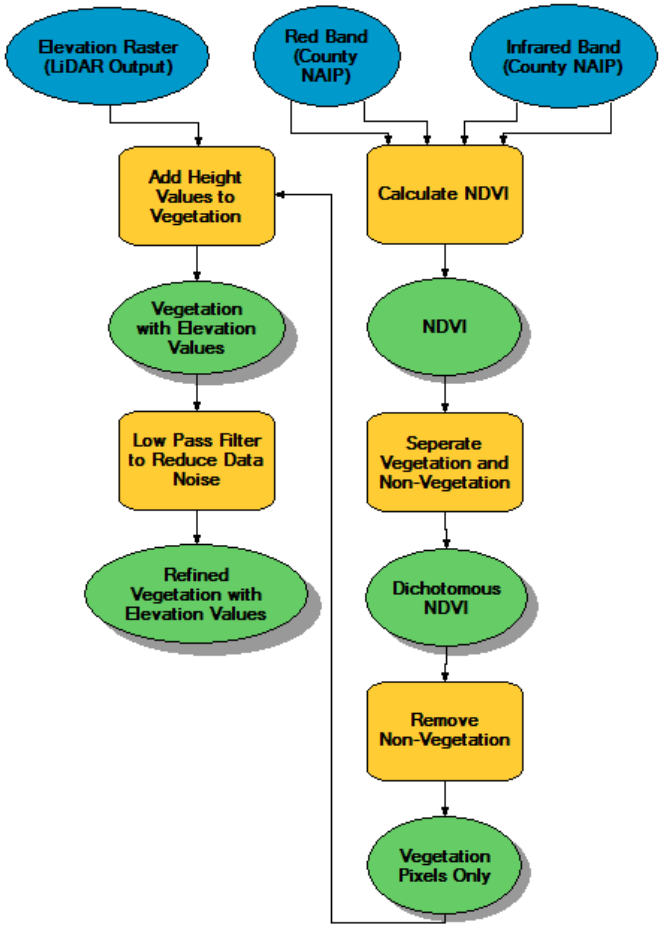

Figure 3.2: Processing steps for production of vegetation with elevations raster layer. The right column shows the steps for extracting vegetation from the NAIP imagery, and the left column shows the addition of heights to the vegetation pixels.

2.3 *LiDAR Contouring and Depression Identification*

Another technique for identifying potential wetlands locations is to find low lying areas. This will not be able to identify all kinds of wetlands due to their great variability, but may be helpful for spotting those locations with greater potential to be wet areas because they are at the bottom of a depression. LiDAR

files can be used to create a number elevation products. In the previous section I created DEMs, DSMs, and combined them to get a heights raster. In this section I used the same LiDAR points, but processed them in a new way. Instead of creating a raster elevation file, the LiDAR data was processed into bare earth contour lines. These lines were then compared to each other to determine which lines were located at the bottom of a depression. The technique was developed by Wu et al. (2015), in which depressions were identified using contour lines generated from smoothed DEMs. After trial and error in ArcMap, I developed a similar process using ArcMap tools to process LiDAR files. Tool usage and settings were determined by evaluating which outputs were more useful for depression identification. I created a Python script tool for this workflow (Appendix 3), an overview of which is as follows: (1) create contours from a single LiDAR file, (2) convert these contours into polygons, (3) generate statistics for the neighbors of each polygon, and (4) use these statistics to identify which polygons are the center of concentric depressions. There were several intermediate files which could have been optionally output by the script tool, but are not necessary, some of which are contour lines, closed loop polygons, and statistics tables (Figure 3.4). This script can be easily run on multiple files using the iteration loop in the beginning of the LiDAR model used to create elevation rasters in the previous section (Figure 3.1). As this study was largely exploratory, this product was not formally evaluated for accuracy via field or remote methods, and instead is intended for use as a visual aid to identify potential wetlands depression centers.

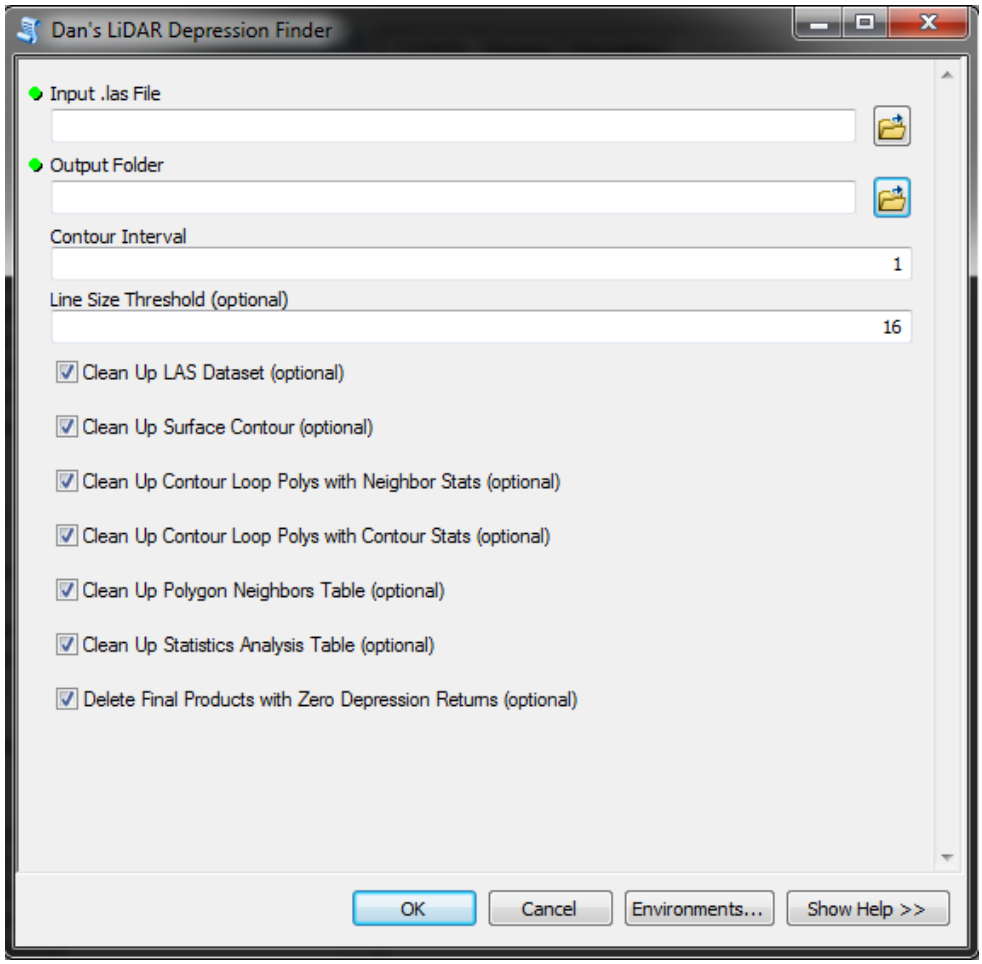

Figure 3.3: Dialog box of the script tool created for isolating depressions from a single LiDAR file. See Appendix 3 for full script

2.4 *Ancillary Data Products for Visualization*

There are a number of additional data products that can be useful for visualizing wetlands locations. I have chosen two of these to be included in my deliverable package as reference materials, namely the National Wetlands Inventory (NWI), and NRCS Soil Database SSURGO. These products will not be used for analysis, and instead will serve as visualization aids. The NWI dataset, produced by the U.S. Fish and Wildlife Service, is a widely-used data product for identifying wetland locations, and has been used by a number of studies (e.g., Dvorett et al. 2016; Enwright et al. 2011; Martin et al. 2012). Production of the NWI dataset began in the 1974 following the Cowardin et al. (1979) classification system. Since then the dataset has been periodically updated, but continues to rely on manual interpretation of aerial

photographs by image analysts as the primary form of wetlands and deepwater habitat detection (U.S. Fish and Wildlife Service 2017). The NRCS Soil Database polygons have also been included for reference because they can be used to identify possible hydric soil conditions, and indicate soil wetness throughout the year. This dataset is produced by the United States Department of Agriculture - National Resources Conservation Service (USDA – NRCS), using field methods of soil type determination and laboratory sample analysis (Soil Survey Staff 2017).

3. Results

3.1 *Vegetation with Heights*

This product combines 4-band NAIP imagery and LiDAR elevations to produce an output containing extracted vegetative surfaces with heights (Figure 3.5). One meter resolution outputs for each county took 12-15 hours to process from start to finish. The output was evaluated for accuracy following Congalton (1991), and has an estimated overall accuracy of 89% for distinguishing trees from groundcover and non-vegetated surfaces. Intermediate products included NDVI, DEMs, DSMs, and elevation rasters (Figure 3.5).

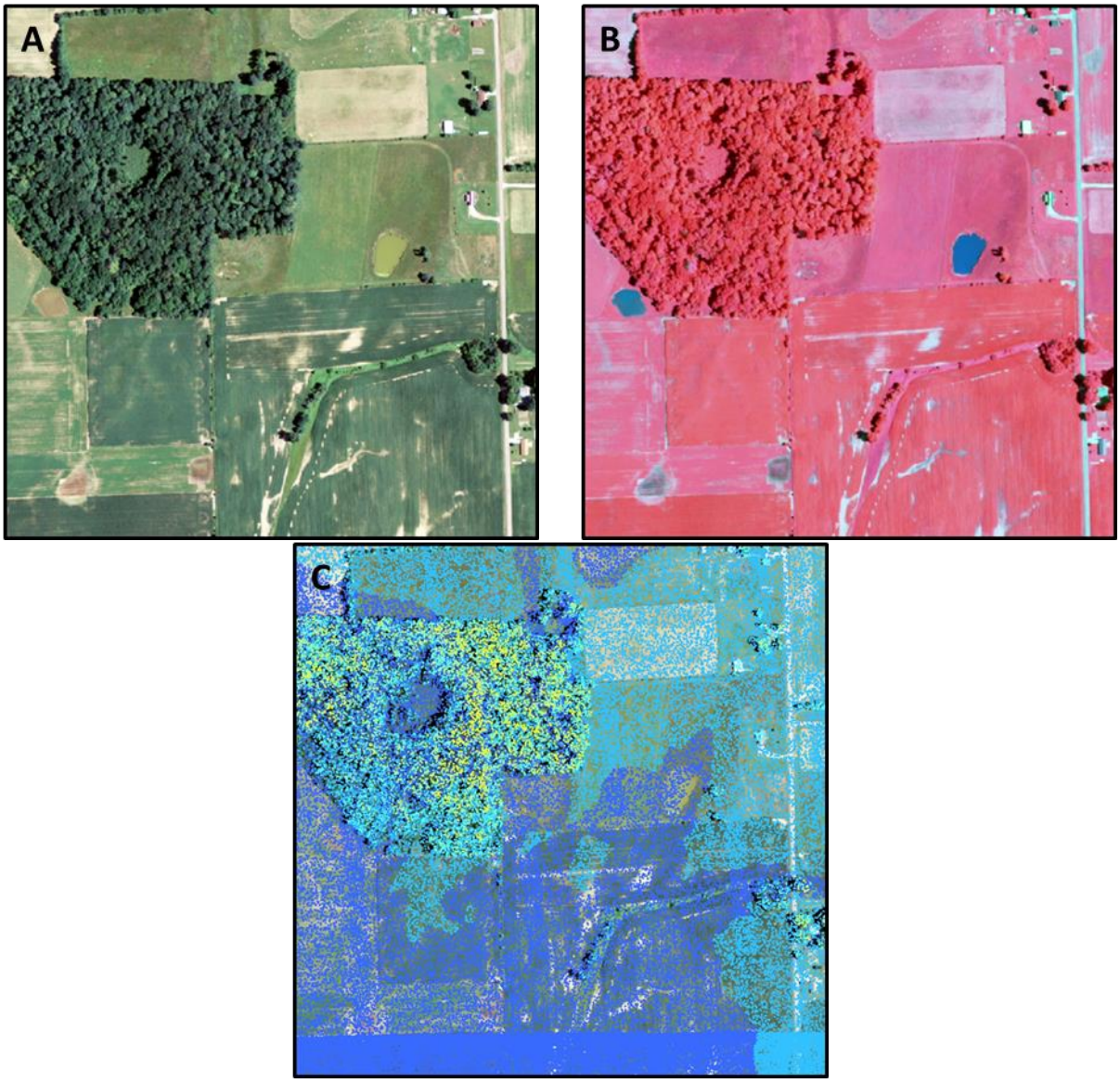

Figure 3.4: Input imagery used to create the vegetation with heights output. (**A**) True color NAIP Imagery with 1m resolution. (**B**) False color infrared NAIP Imagery. The vegetation is vibrantly red in the image because of its high reflectance of infrared radiation. (**C**) LiDAR point cloud data. Each geographically referenced point contains an elevation and a return class, which can be used to differentiate points which hit the ground from points which hit the tops of trees or buildings. Raster symbology displayed with ArcMap defaults except where specified. Area shown is in Henry County, IN, at 39.83, -85.47 decimal degrees

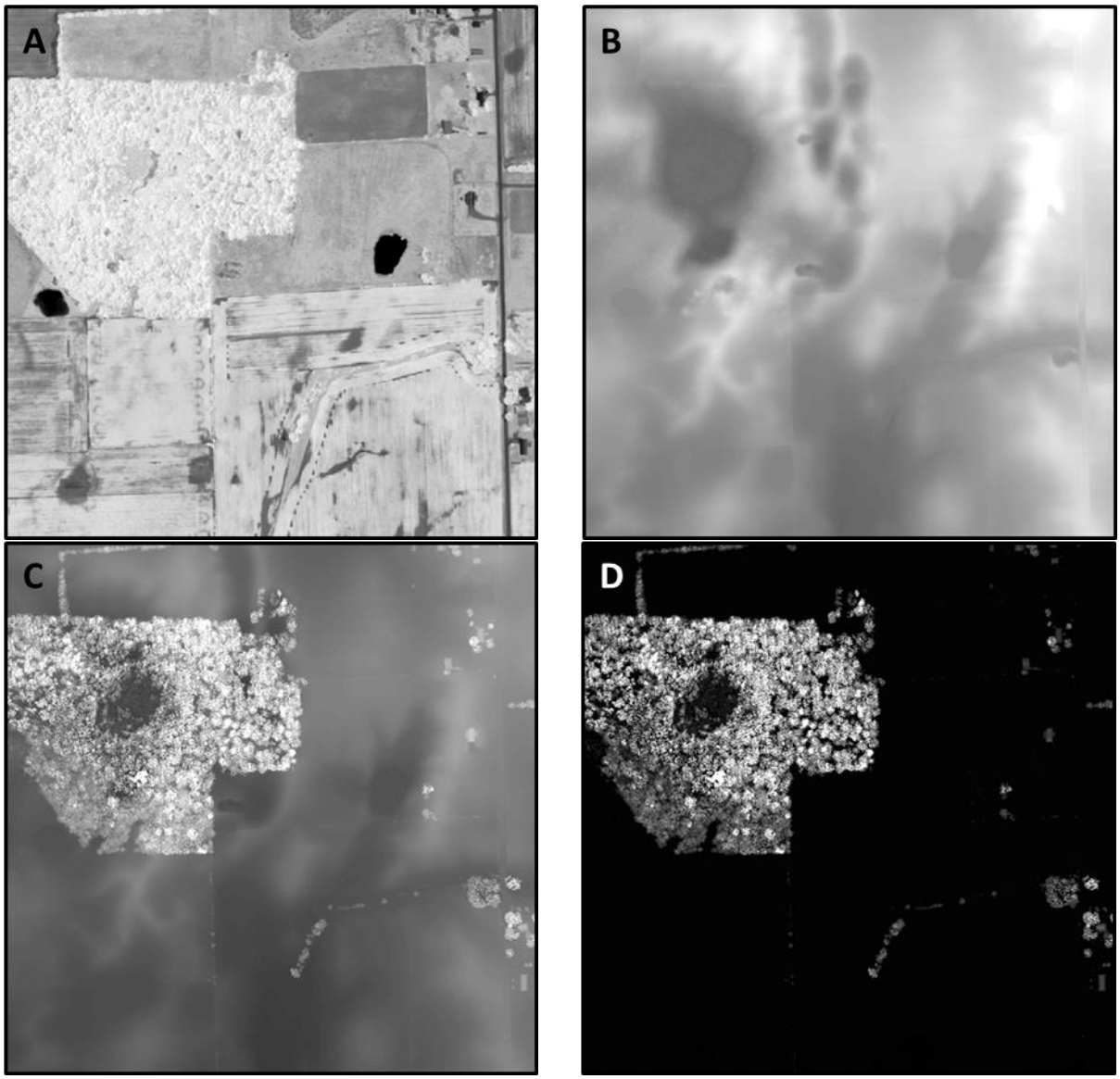

Figure 3.5: Intermediate processing products for the creation of the vegetation with heights output. (**A**) NDVI created from NAIP imagery. Pixels which are more likely to be vigorous vegetation show up lighter in this image, while open water appears black. (**B**) LiDAR generated DEM created using only ground return points. This shows the elevation of the bare earth. Lower areas are darker in color and higher areas are lighter. (**C**) LiDAR generated DSM created using all return points. This contains the elevation of the tops of trees, buildings, and crops where they occur. Lower areas are darker in color and higher areas are lighter. (**D**) Height raster calculated from DEM and DSM products. This raster contains heights values for the pixel returns that are comparable to the actual heights of the object. Taller objects are indicated by lighter shades. Raster symbology displayed with ArcMap defaults except where specified. Area shown is in Henry County, IN, at 39.83, -85.47 decimal degrees.

Figure 3.6: Vegetation with heights output. This product is created by combining vegetation extracted from the NDVI with the height raster. It shows only those pixels which are likely to be vegetation, along with their heights, and can be used to separate vegetation classes based on height. In this image, the vegetation with heights output is semi-transparent and overlain on the true color NAIP imagery. Darker green represents taller vegetation. This product was generated with 1m resolution. Area shown is in Henry County, IN, at 39.83, -85.47 decimal degrees.

3.2 *LiDAR Depression Finder Script Tool*

This product used LiDAR data to create contour lines and search for closed loop depression centers (Figure 3.6). For Delaware County, the script tool generated 5,574,284 contour lines from the 533 individual LiDAR files, and identified 128,814 potential depressions. For Henry County, the tool generated 7,086,495 contour lines and identified 14,627 potential depressions from the 527 Henry LiDAR files. This process took around 5 hours for each county. The resulting depression polygons were visually reviewed to ensure that outputs were consistent and matched study goals; however, no statistical evaluation was conducted.

Figure 3.7: Surface contours and final output from the LiDAR depression finder script. (**A**) LiDAR generated surface contours with 1 ft intervals. Warmer colors have higher elevations. (**B**) LiDAR depression finder script output. Solid white lines indicate potential depression centers. Dashed white lines are NWI polygons for reference. Area shown is in Henry County, IN, at 39.83, -85.47 decimal degrees.

4. Discussion

The data products created in this exploratory study were high resolution, large datasets which contain a lot of information, but their practical usefulness for wetlands identification is uncertain. The vegetation with heights output was relatively accurate, and great for distinguishing tall vegetation like trees from shorter groundcovers, but this information is not specific to wetlands. I did find it useful for identifying areas of mid height vegetation, such as shrubs, within a larger forested region, which can be indicative of certain types of wetlands including buttonbush swamps or wet fields full of hydrophytic grasses. One such area can be seen in the middle of the forested region on the vegetation with heights example image, and is part of an NWI polygon (Figure 3.5, 3.6). One issue that I had with this product was the artificial reduction of heights by the low pass filter, which meant that tree heights were likely underestimated compared to actual tree heights. This filtering was necessary to create more accurate representations of trees because of the way that the LiDAR intersected with the trees. The LiDAR does not hit just the very top of the trees but intersects foliage throughout, making trees look splotchy in the original DSM images (Figure 3.5). I may have been able to remedy this issue by experimenting with different LiDAR sampling and filtering options, but the issue may be more of a nuisance than an actual problem for practical use of the data. This product could potentially have been more suited to wetlands detection if the NAIP imagery had contained a fifth middle infrared band. I would have to been able to create an output containing vegetation wetness which may have significantly improved my results by allowing me to differentiate wet areas. I did explore using the normalized difference water index (NDWI) of McFeeters (1996) which uses only the red and infrared bands that were available, however, after visual examination I found the index to be very similar to the NDVI output, but with opposite values (e.g. water yields low values in NDVI, and high values in NDWI).

I did explore other methods of locating possible wetland areas. I examined several other vegetation indices including simple ratio, differential vegetation index, soil adjusted vegetation index,

and optimized soil adjusted vegetation index, but did not retain these products for further analysis because they did not appear to perform better than NDVI. Another method I explored was to create an atypical 5-band NAIP image by combining the NDVI back in to the NAIP image as the fifth band, and using this product along with known wetland locations for the maximum likelihood classification method available in ArcMap, but the results were unsuitable for further analysis due to poor classification performance. I also explored many different ways of creating and processing contours for my script tool. I read in multiple studies that common practice was to create contours from a smoothed DEM, but I experienced technical limitations due to the large volumes of data I was processing.

I think my script tool worked very well for pulling out a specific type of closed loop depression center contour from all other closed loops. The tool successfully finds the bottom-most ring of a concentric depression and isolates it for further examination. After visually inspecting the results I found that the majority of the output polygons appear to fall into two classes, actual depression centers and artifacts. All of the artifacts that I examined were the result of polygons looping back onto themselves, and therefore this problem has a number of solutions which could have been implemented. One solution would be to create topology rules in ArcMap (e.g. lines cannot touch each other, and cannot loop back upon themselves) and solve the corresponding errors either manually or through considerable automation for each file as it processes; I suspect that both ways of implementing this solution would drastically increase processing time. Another solution which could reduce artifacts would be to use smoothed DEMs instead of processing directly from LiDAR as I did, but the additional processing steps required to create smoothed DEMs from the LiDAR would also significantly increase processing time. I have also read that there is a tool in the ArcGIS Production Mapping extension, which was unavailable to me, called Remove Self Intersections, that is specific to finding recursive polygon segments, which also may have been able to help with this issue (ESRI 2017).

Another caveat of this script tool is that some depressions cannot be found due to a strict set of rules I implemented for filtering outputs (Appendix 3; line 160 – 165). While I experimented with a number of other rule sets, the selected technique produced the most positive outputs while reducing as many artifacts and false positives as possible. Of the false positives, the most difficult to filter were hill complexes with multiple high points. The lower of the multiple high points would often be mistaken for a depression because of its relationship with the contour line below it, and that line's relationship with the higher hilltop. To eliminate this problem, I had to implement the rule on line 164 (Appendix 3), which in addition to preventing false positives from hill complexes, also eliminated locations with depressions that are not part of a concentric arrangement (e.g. depressions along the side of a hill, and multiple pits in the center of a depression). Crafting a solution to this shortcoming could substantially improve the tool's performance for identifying hillside depressions, which may contain ecologically valuable wetlands.

Lastly, I ran into a few issues with the ArcGIS tools that I used, which could have made the process much simpler. The primary issue was a documented problem with the Feature to Polygons tool. When creating polygons from the contours, the contour value is not carried over, and the polygons do not contain values. To fix this I had to spatially join the contour line values back onto the polygons, but this is also not as simple as it sounds, because, by the nature of the creating polygons from contours, each contour line touches a minimum of two different polygons which can be their spatial join match. My solution to this was to run though many statistics options during the spatial join, so that I had more contour line numbers to work with. However, both of these problems become more complicated for non-concentric topographic arrangements, which probably accounts for some of my difficulty in separating depressions from hills with multiple tops and other varied landforms. If this problem with the Feature to Polygons tool is remedied within ArcGIS, my script could be simplified and may be more successful at identifying true landscape depressions.

Overall this script does a pretty good job of locating potential depression centers which can later be examined on a case by case basis for wetlands conservation efforts. Improvements could be made by developing a better set of depression filtering rules which allowed for greater variability in depression types. The NAIP imagery and LiDAR vegetation with heights output are great resources for visualizing vegetation, but may not be directly applicable to detecting all types of wetlands. Along with NWI polygons and soil hydrography data, these products may provide more information to RCAS than their previously available conservation map. Unfortunately, because of the sheer size of the data products which I have created, online map hosting is not a feasible data visualization outcome. I would have liked to deliver the products within a polished online mapping application, but instead, these data products are being delivered to RCAS personnel on a portable hard drive containing premade maps for ArcGIS Explorer, a free software package for visualizing GIS data. The two main products I have created are two of many possible outcomes, and I believe that there are more methodologies that could be explored related to wetlands mapping. However, with the currently available free data products and challenges with distinguishing wetlands from other land covers, wetlands detection is a rather difficult task.

5. Conclusion

As more and more data products become freely available at the state, county, or local level, conducting high resolution analysis becomes more feasible. Wetlands identification has been and continues to be a challenging topic because of the large variability of topographic, morphologic, and vegetative features of which they are comprised, and because of their ephemeral nature. This study generated two high resolution outputs—vegetation with heights and landscape depressions—which could be useful for identifying potentially valuable wetlands. These products were delivered to RCAS for use as references, so that they can better focus conservation efforts in East Central Indiana.

CITATIONS

Aerial Photography Field Office. 2008. Qualitative and Quantitative Synopsis on NAIP Usage from 2004-2008. In: USDA FSA, editor. Salt Lake City, UT.

Alonzo M, McFadden JP, Nowak DJ, Roberts DA. 2016. Mapping urban forest structure and function using hyperspectral imagery and lidar data. Urban Forestry & Urban Greening 17:135-147.

Anguelov D, Dulong C, Filip D, Frueh C, Lafon S, Lyon R, Ogale A, Vincent L, Weaver J. 2010. Google Street View: Capturing the World at Street Level. IEEE Computer Society.

Armson D, Stringer P, Ennos AR. 2013. The effect of street trees and amenity grass on urban surface water runoff in Manchester, UK. Urban Forestry & Urban Greening 12:282-286.

Badland HM, Opit S, Witten K, Kearns RA, Mavoa S. 2010. Can virtual streetscape audits reliably replace physical streetscape audits? J Urban Health 87(6):1007-16.

Baró F, Chapparro L, Gómez-Baggethun E, Langemeyer J, Nowak DJ, Terradas J. 2014. Contribution of Ecosystem Services to Air Quality and Climate Change Mitigation Policies: The Case of Urban Forests in Barcelona, Spain. Ambio 43:466-479.

Bartens J, Day SD, Harris JR, Dove JE, Wynn TM. 2008. Can urban tree roots improve infiltration through compacted subsoils for stormwater management? Journal of Environmental Quality 37:2048-2057.

Baumgardner D, Varela S, Escobedo FJ, Chacalo A, Ochoa C. 2012. The role of a peri-urban forest on air quality improvement in the Mexico City megalopolis. Environmental Pollution 163:174-183.

Berland A. 2012. Long-term urbanization effects on tree canopy cover along an urban–rural gradient. Urban Ecosystems 15:721-738.

Berland A, Hopton ME. 2014. Comparing street tree assemblages and associated stormwater benefits among communities in metropolitan Cincinnati, Ohio, USA. Urban Forestry & Urban Greening 13(4):734-741.

Berland A, Lange DA. 2017. Google Street View shows promise for virtual street tree surveys. Urban Forestry & Urban Greening.

Berland A, Shiflett SA, Shuster WD, Garmestani AS, Goddard HC, Herrmann DL, Hopton ME. 2017. The role of trees in urban stormwater management. Landscape and Urban Planning 162:167-177.

Bolund P, Hunhammar S. 1999. Ecosystem services in urban areas. Ecological Economics 29:293-301.

Clarke P, Ailshire J, Melendez R, Bader M, Morenoff J. 2010. Using Google Earth to conduct a neighborhood audit: reliability of a virtual audit instrument. Health Place 16(6):1224-9.

Congalton RG. 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment 37:35-46.

Cowardin LM, Carter V, Golet FC, Laroe ET. 1979. Classification of Wetlands and Deepwater Habitats of the United States. Washington D.C.: U.S. Fish and Wildlife Service.

Cowett FD. 2012. Modeling street trees on a statewide basis in New York State. ProQuest Dissertations Publishing.

Cowett FD. 2014. Methodology for Spatial Analysis of Municipal Street Tree Benefits. Arboriculture & Urban Forestry 40(2):112-118.

Dahl TE. 2000. Status and Trends Wetlands in the Conterminous United States 1986 to 1997. Washington DC: U.S. Fish and Wildlife Service.

Dahl TE. 2011. Status and Trends of Wetlands in the Conterminous United States 2004 to 2009. Washington DC: U.S. Fish and Wildlife Service.

Davranche A, Lefebvre G, Poulin B. 2010. Wetland monitoring using classification trees and SPOT-5 seasonal time series. Remote Sensing of Environment 114(3):552-562.

Denman L. Are street trees and their soils an effective stormwater treatment measure? 7th National Street Tree Symposium; 2006; Adelaide, Australia.

Donovan GH, Butry DT. 2010. Trees in the city: valuing street trees in Portland, Oregon. Landscape and Urban Planning 94:77-83.

Dvorett D, Davis C, Papes M. 2016. Mapping and Hydrologic Attribution of Temporary Wetlands Using Recurrent Landsat Imagery. Wetlands 36(3):431-443.

Enwright N, Forbes MG, Doyle RD, Hunter B, Forbes W. 2011. Using Geographic Information Systems (GIS) to Inventory Coastal Prairie Wetlands Along the Upper Gulf Coast, Texas. Wetlands 31(4):687-697.

Escobedo FJ, Adams DC, Timilsina N. 2014. Urban forest structure effects on property value. Ecosystem Services 12:209-217.

Escobedo FJ, Kroeger T, Wagner JE. 2011. Urban forests and pollution mitigation: Analyzing ecosystem services and disservices. Environmental Pollution 159:2078-2087.

ESRI. Remove Self Intersections [Internet]. Available from: http://desktop.arcgis.com/en/arcmap/latest/tools/production-mapping-toolbox/remove-self-intersections.htm#L_

Fazio JR. 2010. How Trees Can Retain Stormwater Runoff. Tree City USA Bulletin 55.

Fischer BC, Steed BC. 2008. Street trees - a misunderstood common-pool resource. 84th International Society of Arboriculture Annual Conference. St. Louis, MO. p. 18.

Frome A, Cheung G, Abdulkader A, Zennaro M, Bo W, Bissacco A, Adam H, Neven H, Vincent L. 2009. Large-scale privacy protection in Google Street View.2373-2380.

Gallant AL. 2015. The Challenges of Remote Monitoring of Wetlands. Remote Sensing 7(8):10938-10950.

Gorman J. 2004. Residents' opinions on the value of street trees depending on tree location. Journal of Arboriculture 30(1):36-44.

Hardin P, Jensen R. 2007. The effect of urban leaf area on summertime urban surface kinetic temperatures: A Terre Haute case study. Urban Forestry & Urban Greening 6:63-72.

Huete AR. 1988. A soil-adjusted vegetation index (SAVI). Remote Sens Environ 25:295-309.

i-Tree. Reference Cities—The Science Behind i-Tree Streets (STRATUM) [Internet]. Available from: https://www.itreetools.org/streets/resources/Streets_Reference_Cities_Science_Update_Nov2011.pdf

ISDP. 2011-2013 Indiana Orthophotography (RGBI), LiDAR and Elevation [Internet]. Available from: http://gis.iu.edu/datasetInfo/statewide/in_2011.php

Johnson WC, Millet BV, Gilmanov T, Voldseth RA, Guntenspergen GR, Naugle DE. 2005. Vulnerability of northern prairie wetlands to climate change. BioScience 55:863-872.

Kardan O, Gozdyra P, Misic B, Moola F, Palmer LJ, Paus T, Berman MG. 2015. Neighborhood greenspace and health in a large urban center. Scientific Reports 5:11610.

Ko Y, Lee J-H, McPherson EG, Roman LA. 2015. Long-term monitoring of Sacramento Shade program trees: Tree survival, growth and energy-saving performance. Landscape and Urban Planning 143:183-191.

Kondo MC, Low SC, Henning J, Branas CC. 2015. The impact of green stormwater infrastructure installation on surrounding health and safety. American Journal of Public Health 105(3):e114-e121.

Kovacs KF, Haight RG, Jung S, Locke DH, O'Neil-Dunne J. 2013. The marginal cost of carbon abatement from planting street trees in New York City. Ecological Economics 95:1-10.

Landry SM, Chakraborty J. 2009. Street trees and equity: evaluating the spatial distribution of an urban amenity. Environment and Planning A 41:2651-2670.

Less EL, McKee P, Toomey T, Nelson T, Erickson D, Xiong S, Jones-Webb R. 2015. Matching study areas using Google Street View: A new application for an emerging technology. Eval Program Plann 53:72-9.

Levick SR, Rogers KH. LiDAR and Object-Based Image Analysis as Tools for Monitoring the Structural Diversity of Savanna Vegetation. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences; 2006. p. 6.

Li X, Zhang C, Li W, Kuzovkina YA, Weiner D. 2015a. Who lives in greener neighborhoods? The distribution of street greenery and its association with residents' socioeconomic conditions in Hartford, Connecticut, USA. Urban Forestry & Urban Greening 14(4):751-759.

Li X, Zhang C, Li W, Ricard R, Meng Q, Zhang W. 2015b. Assessing street-level urban greenery using Google Street View and a modified green view index. Urban Forestry & Urban Greening 14(3):675-685.

Livesley SJ, Baudinette B, Glover D. 2014. Rainfall interception and stem flow by eucalypt street trees – the impacts of canopy density and bark type. Urban Forestry & Urban Greening 13(1):192-197.

Long Y, Liu L. 2017. How green are the streets? An analysis for central areas of Chinese cities using Tencent Street View. PLoS One 12(2):e0171110.

Lovasi GS, Quinn JW, Neckerman KM, Perzanowski MS, Rundle A. 2008. Children living in areas with more street trees have lower prevalence of asthma. Journal of Epidemiology and Community Health 62:647-649.

Martin GI, Kirkman LK, Hepinstall-Cymerman J. 2012. Mapping Geographically Isolated Wetlands in the Dougherty Plain, Georgia, USA. Wetlands 32(1):149-160.

Martin NA, Chappelka AH, Keever GJ, Loewenstein EF. 2011. A 100% Tree Inventory Using i-Tree Eco Protocol: A Case Study at Auburn University, Alabama, U.S. Arboriculture & Urban Forestry 37(5):207-212.

Martin NA, Chappelka AH, Somers G, Loewenstein EF, Keever GJ. 2013. Evaluation of Sampling Protocol for i-Tree Eco: A Case Study in Predicting Ecosystem Services at Auburn University. Arboriculture & Urban Forestry 39(2):56-61.

McFeeters SK. 1996. The use of the Normalised Difference Water Index (NDWI) in the delineation of open water features. International Journal of Remote Sensing 17:1425-1432.

McPherson EG. 2010. Selecting reference cities for i-Tree Streets. Arboriculture & Urban Forestry 36(5):230-240.

McPherson EG, Kotow L. 2013. A municipal forest report card: results for California, USA. Urban Forestry & Urban Greening 12:134-143.

McPherson EG, Nowak D, Heisler G, Grimmon S, Souch C, Grant R, Rowntree R. 1997. Quantifying urban forest structure, function, and value: the Chicago Urban Forest Climate Project. Urban Ecosystems 1:49-61.

McPherson EG, Rowntree RA. 1989. Using structural measures to compare twenty-two U.S. street tree populations. Landscape Journal 8(1):13-23.

McPherson EG, Xiao Q, Aguaron E. 2013. A new approach to quantify and map carbon stored, sequestered and emissions avoided by urban forests. Landscape and Urban Planning 120:70-84.

McPherson G, Simpson JR, Peper PJ, Maco SE, Xiao Q. 2005. Municipal Forest Benefits and Costs in Five US Cities. Journal of Forestry 103(8):411-416.

Millward AA, Sabir S. 2011. Benefits of a forested urban park: What is the value of Allan Gardens to the city of Toronto, Canada? Landscape and Urban Planning 100(3):177-188.

Millward AA, Torchia M, Laursen AE, Rothman LD. 2014. Vegetation Placement for Summer Built Surface Temperature Moderation in an Urban Microclimate. Environmental Management 53(6):1043-1057.

Mullaney J, Lucke T, Trueman SJ. 2015. A review of benefits and challenges in growing street trees in paved urban environments. Landscape and Urban Planning 134:157-166.

Nowak DJ, Crane DE, Stevens JC. 2006a. Air pollution removal by urban trees and shrubs in the United States. Urban Forestry & Urban Greening 4(3-4):115-123.

Nowak DJ, Greenfield EJ, Hoehn RE, Lapoint E. 2013a. Carbon storage and sequestration by trees in urban and community areas of the United States. Environmental Pollution 178:229-236.
Nowak DJ, Hirabayashi S, Bodine A, Hoehn R. 2013b. Modeled PM2.5 removal by trees in ten U.S. cities and associated health effects. Environmental Pollution 178:395-402.
Nowak DJ, Hoehn R. 2007. Oxygen Production by Urban Trees in the United States. Arboriculture & Urban Forestry 33(3):220–226.
Nowak DJ, Hoehn RE, Bodine AR, Greenfield EJ, O'Neil-Dunne J. 2013c. Urban forest structure, ecosystem services and change in Syracuse, NY. Urban Ecosystems 19(4):1455-1477.
Nowak DJ, Hoehn RE, III, Crane DE, Stevens JC, Walton JT, Bond J, Ina G. 2006b. Assessing urban forest effects and values: Minneapolis' urban forest. Newtown Square, PA: USDA Forest Service.
O'Neil-Dunne J, MacFaden S, Royar A. 2014. A versatile, production-oriented approach to high-resolution tree-canopy mapping in urban and suburban landscapes using GEOBIA and data fusion. Remote Sensing 6(12):12837-12865.
Ozesmi SL, Bauer ME. 2002. Satellite remote sensing of wetlands. Wetlands Ecology and Management 10:381–402.
Pearson RL, Miller LD. Remote mapping of standing crop biomass for estimation of the productivity of the short-grass Prairie, Pawnee National Grasslands, Colorado. 8th International Symposium on Remote Sensing of the Environment; 1972; Ann Arbor, MI: Environmental Research Institute of Michigan. p. 1357-1381.
Rakower LH. 2011. Blurred Line: Zooming in on Google Street View and the Global Right to Privacy. Brooklyn Journal of International Law 37:317-347.
RCAS. Chapter History [Internet]. Available from: http://cooperaudubon.org/chapter_history.html
Richards DR, Edwards PJ. 2017. Quantifying street tree regulating ecosystem services using Google Street View. Ecological Indicators 77:31-40.
Richardson AJ, Everitt JH. 1992. Using spectral vegetation indices to estimate rangeland productivity. Geocarto International 7(1):63-69.
Roman A, Garg G, Levoy M. Interactive Design of Multi-Perspective Images for Visualizing Urban Landscapes. Visualization; 2004.
Roman A, Lensch HPA. Automatic Multiperspective Images. In: Akenine-Moller T, Heidrich W, editors. Eurographics Symposium on Rendering; 2006.
Roman LA, Scharenbroch BC, Östberg JPA, Mueller LS, Henning JG, Koeser AK, Sanders JR, Betz DR, Jordan RC. 2017. Data quality in citizen science urban tree inventories. Urban Forestry & Urban Greening 22:124-135.
Rondeaux G, Steven M, Baret F. 1996. Optimization of Soil-Adjusted Vegetation Indices. Remote Sens Environ 55:95-107.
Rouse JW, Haas RH, Schell JA, Deering DW. 1973. Monitoring vegetation systems in the great plains with ERTS. Third ERTS Symposium. NASA. p. 309-317.
Rousselet J, Imbert CE, Dekri A, Garcia J, Goussard F, al. e. 2013. Assessing Species Distribution Using Google Street View: A Pilot Study with the Pine Processionary Moth. PLoS ONE 8(10).
Sander H, Polasky S, Haight RG. 2010. The value of urban tree cover: A hedonic property price model in Ramsey and Dakota Counties, Minnesota, USA. Ecological Economics 69:1646-1656.
Sanders RA. 1981. Diversity in the street trees of Syracuse, New York. Urban Ecology 5:33-43.
Schroeder H, Flannigan J, Coles R. 2006. Residents' attitudes toward street trees in the UK and U.S. communities. Arboriculture & Urban Forestry 32(5):236-246.
Schroeder HW, Cannon WN, Jr. 1983. The Esthetic Contribution of Trees to Residential Streets in Ohio Towns. Journal of Arboriculture 9(9):237-243.
Seiferling I, Naik N, Ratti C, Proulx R. 2017. Green streets − quantifying and mapping urban trees with street-level imagery and computer vision. Landscape and Urban Planning 165:93-101.

Siegert NW, McCullough DG, Liebhold AM, Telewski FW, MacIsaac H. 2014. Dendrochronological reconstruction of the epicentre and early spread of emerald ash borer in North America. Diversity and Distributions 20(7):847-858.

Soares AL, Rego FC, McPherson EG, Simpson JR, Peper PJ, Xiao Q. 2011. Benefits and costs of street trees in Lisbon, Portugal. Urban Forestry & Urban Greening 10(2):69-78.

Soil Survey Staff. Description of SSURGO Database [Internet]. Natural Resources Conservation Service, United States Department of Agriculture. Available from: https://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/survey/geo/?cid=nrcs142p2_053627

Stovin VR, Jorgensen A, Clayden A. 2008. Street trees and stormwater management. Arboricultural Journal 30:297-310.

Tsai VJD, Chang C-T. 2013. Three-dimensional positioning from Google street view panoramas. IET Image Processing 7(3):229-239.

Tsunetsugu Y, Lee J, Park B-J, Tyrväinen L, Kagawa T, Miyazaki Y. 2013. Physiological and psychological effects of viewing urban forest landscapes assessed by multiple measurements. Landscape and Urban Planning 113:90-93.

Tzoulas K, Korpela K, Venn S, Yli-Pelkonen V, Kaźmierczak A, Niemela J, James P. 2007. Promoting ecosystem and human health in urban areas using Green Infrastructure: A literature review. Landscape and Urban Planning 81(3):167-178.

U.S. Fish and Wildlife Service. 2002. National wetlands inventory: A strategy for the 21st century. Washington D.C.: U.S. Fish and Wildlife Service.

U.S. Fish and Wildlife Service. NWI Overview [Internet]. Available from: https://www.fws.gov/Wetlands/nwi/Overview.html

USDA's FSA. NAIP Imagery [Internet]. Available from: https://www.fsa.usda.gov/programs-and-services/aerial-photography/imagery-programs/naip-imagery/

Wang Y. 2011. Remote sensing of protected lands. CRC Press.

Wang Y, Bakker F, de Groot R, Wörtche H. 2014. Effect of ecosystem services provided by urban green infrastructure on indoor environment: A literature review. Building and Environment 77:88-100.

Wegner J, Branson S, Hall D, Schindler K, Perona P. 2016. Cataloging Public Objects Using Aerial and Street-Level Images - Urban Trees. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas.

Wolf KL. 2003. Public Response to the Urban Forest in Inner-City Buisness Districts. Journal of Aboriculture 29(3):10.

Wu Q, Liu H, Wang S, Yu B, Beck R, Hinkel K. 2015. A localized contour tree method for deriving geometric and topological properties of complex surface depressions based on high-resolution topographical data. International Journal of Geographical Information Science 29(12):2041-2060.

Wu QS, Lane CR. 2016. Delineation and Quantification of Wetland Depressions in the Prairie Pothole Region of North Dakota. Wetlands 36(2):215-227.

Wu YT, Nash P, Barnes LE, Minett T, Matthews FE, Jones A, Brayne C. 2014. Assessing environmental features related to mental health: a reliability study of visual streetscape images. BMC Public Health 14:1094.

Xiao Q, McPherson EG. 2002. Rainfall interception by Santa Monica's municipal urban forest. Urban Ecosystems 6:291-302.

Xiao Q, McPherson EG, Simpson JR, Ustin SL. 1998. Rainfall interception by Sacramento’s urban forest. Journal of Arboriculture 24(4):235-244.

Xiao Q, McPherson EG, Ustin SL, Grismer ME, Simpson JR. 2000. Winter rainfall interception by two mature open-grown trees in Davis, California. Hydrological Processes 14:763-784.

Yan WY, Shaker A, Easa S. 2013. Potential Accuracy of Traffic Signs' Positions Extracted From Google Street View. IEEE Transactions on Intelligent Transportation Systems 14(2):1011-1016.

APPENDIX 1: GOOGLE STREET VIEW SHOWS PROMISE FOR VIRTUAL STREET TREE SURVEYS – AUTHOR'S PREPRESS VERSION

Google Street View shows promise for virtual street tree surveys

Adam Berland & Daniel A. Lange

Department of Geography, Ball State University, Muncie, IN, USA

Abstract

Geospatial technologies are increasingly relevant to urban forestry, but their use may be limited by cost and technical expertise. Technologies like Google Street View™ are appealing because they are free and easy to use. We used Street View to conduct a virtual survey of street trees in three municipalities, and compared our results to existing field data from the same locations. The virtual survey analyst recorded the locations of street trees, identified trees to the species level, and estimated diameter at breast height. Over 93% of the 597 trees documented in the field survey were also observed in the virtual survey. Tree identification in the virtual survey agreed with the field data for 90% of trees at the genus level and 66% of trees at the species level. Identification was less reliable for small trees, rare taxa, and for trees with multiple species in the same genus. In general, tree diameter was underestimated in the virtual survey, but estimates improved as the analyst became more experienced. This study is the first to report on manual interpretation of street tree characteristics using Street View. Our results suggest that virtual surveys in Street View may be suitable for generating some types of street tree data or updating existing data sets more efficiently than field surveys.

Keywords: municipal forestry; species identification; street-level photography; tree inventories

Citation: Berland A, Lange DA. 2017. Google Street View shows promise for virtual street tree surveys. *Urban Forestry & Urban Greening* 21: 11-15. doi: 10.1016/j.ufug.2016.11.006.

PREPRESS VERSION

The final publication is available at http://dx.doi.org/10.1016/j.ufug.2016.11.006.

Introduction

Street trees—trees in the public right-of-way along streets—are a prominent component of the urban forest (Berland and Hopton, 2014; McPherson et al., 2016). Street trees provide valuable environmental, social, and economic benefits (Mullaney et al., 2015). Given the importance of street trees, field-based tree surveys are carried out to facilitate informed management of this municipal resource. Street tree surveys are used to characterize attributes of street tree assemblages such as the number of trees, tree sizes, and species composition (Pedlar et al., 2013). These surveys may be used to identify management needs associated with planting, pruning, tree removal, mitigating hazards and infrastructure conflicts, and understanding the potential and realized impacts of tree pests and pathogens. Furthermore, street tree survey data are frequently used to estimate the environmental and aesthetic benefits of street trees using models like i-Tree Streets (i-Tree, 2016).

Field-based street tree surveys provide invaluable information for municipal forest managers, but they are costly, labor-intensive, time-consuming, and pose safety risks (e.g., automobile crashes) to field crews (Alonzo et al., 2016). These problems can be mitigated through the use of geospatial technologies, which are emerging with broad applications for characterizing and managing the urban forest (Ward and Johnson, 2007). Remote sensing has driven major advances in quantifying urban forest abundance (McGee et al., 2012; O'Neil-Dunne et al., 2014) and quality (Alonzo et al., 2014) using airborne and satellite imagery. Remote sensing techniques have permitted powerful analyses such as early detection of tree pests (San Souci et al., 2009), comparison of tree distributions to population characteristics (Landry and Chakraborty, 2009; Berland et al., 2015), and estimation of tree benefits (Alonzo et al., 2016).

However, despite the capabilities and consistent advances in the use of remote sensing to characterize the urban forest, remote sensing techniques remain largely inaccessible to non-experts including most urban forest managers. In addition, high-quality remote sensing products are generally expensive because they require costly data sets and a paid geospatial analyst. As such, these products are out of reach for most municipalities due to inadequate funding and in-house expertise.

Geospatial technologies that are simple to use and free or inexpensive may be more useful to a larger number of municipal forest managers as compared to sophisticated approaches requiring expensive data. Google Street View™ (GSV) is a geospatial platform that provides ground-based panoramic photographs captured along streets (Anguelov et al., 2010). GSV is available online (http://www.google.com/streetview) or through Google Earth™ (https://www.google.com/earth/), and it provides excellent coverage of streets in the US and many other countries around the world. GSV's street-level perspective of a community is a powerful tool for virtual exploration of neighborhoods, and this technology has been used in several research applications. Li et al. (2015a; 2015b) used GSV to characterize greenery in urban streetscapes by quantifying green pixels in Street View images. Rousselet et al. (2013) used GSV to study the distribution of a forest pest by identifying the pest's distinctive nests. GSV has also been used in public health studies to conduct virtual audits of the built environment (Badland et al., 2010; Clarke et al., 2010). In these diverse applications, GSV has shown promise for providing useful virtual information that reduces the need to visit field sites.

Given the widespread use of GSV, we conducted this exploratory study to understand whether this technology can be employed to conduct robust virtual surveys of street trees. In other words, can we generate reliable data about street trees based solely on GSV imagery? We

assessed the quality of street tree data generated using GSV imagery by comparing it to existing field data. Specifically, we evaluated data regarding the numbers, sizes, and species of trees. As compared to field surveys, we expected that this virtual approach could save time, money, and reduce safety hazards associated with field work. As this study is the first of its kind, this article provides valuable information about the capabilities, limitations, and possible applications of GSV image interpretation in urban forestry research and practice.

Methods

Study area

This study was conducted in the following three municipalities in suburban Cincinnati, OH, USA: Mt. Healthy (39.23° N, 84.55° W; population 6,061), Reading (39.23° N, 84.44° W; population 10,357), and Wyoming (39.23° N, 84.47° W; population 8,404). These municipalities are primarily comprised of low- to medium-density residential land. In each municipality, over 70% of the houses were built before 1970 (US Census Bureau, 2016). Wyoming has an active planting program for street trees; the other two have existing street trees but no current planting program.

Field data collection

In 2013, we randomly sampled street segments in the study area, surveying >10% of the total length of local, public streets in each municipality. On each street segment, we followed the i-Tree Streets sampling protocol (i-Tree, 2016). We recorded the geographic location of every tree encountered in the public right-of-way, with trees defined as woody vegetation >2.5 cm (1 in.) diameter at breast height (dbh). Trees were identified to the species level, and dbh was measured

to the nearest 0.1 cm. A geographic information system (GIS) was used to plot the locations of trees inventoried in the field.

Virtual survey of street trees

We used GSV to virtually survey the same street segments that were sampled in the field. The virtual survey was conducted from October 2015 to March 2016, roughly 2.5 years after the field survey. The analysis was conducted by a single individual who held a bachelor's degree in biology with a focus on field botany. The analyst had never visited the study area, so he had no prior knowledge of which tree species to expect.

GSV was accessed using Google Earth™ Pro, which was freely available and supported the import of GIS shapefiles to view street segment lines within the GSV interface. This feature was useful for ensuring that the field crew and virtual survey analyst both inventoried the same street segments. Along each street segment, the analyst noted each tree's geographic location, identified the tree to the species level, and manually estimated dbh to the nearest cm based on the tree's appearance in the photograph within the landscape context. The analyst also noted the date of the GSV image. We did not collect data regarding tree health, evidence of pests, or infrastructure conflicts in this exploratory study.

The analyst calibrated his virtual survey estimates at multiple stages of the study. Prior to the virtual survey, 20 trees of various species and size classes were selected in our local community; the analyst estimated the dbh and species of these trees using GSV, and visited them in the field to assess his performance. Then during the virtual survey, the analyst completed one municipality at a time, and assessed his performance in that municipality before moving on to the next. After generating unsatisfactory dbh estimates in Mt. Healthy, the analyst used trees of

known dbh to make a pictorial reference guide to assist in estimating dbh based on a tree's appearance in GSV. We assumed this midstream calibration would lead to varying levels of data quality among the study municipalities, but would improve the overall performance of the virtual survey for the study area as a whole.

Comparison of field data and virtual survey data

We compared the virtual survey to the field data to assess the performance of the virtual survey. We approached this comparison with the assumption that the field data were correct, and that disagreement between the two data sets was introduced by the virtual survey (but see Discussion points below about species identification and temporal mismatch between field and virtual data collection). To compare observations, trees from the two surveys were first matched by geographic location using a spatial join in ArcGIS 10.3 (ESRI, 2014) and subsequent manual adjustments. Trees were matched across the two surveys if they were located at the same street address, similarly sized, and represented the same genus or genera that could reasonably be confused.

To compare tree counts, we noted the total number of trees in each survey, the number of matched trees, and the percent of street segments on which the tree counts agreed between the two surveys. For matched trees, we compared the level of agreement in tree identification and size class estimates between the field and virtual surveys using raw percent agreement and Cohen's kappa. Kappa tests whether the observed level of agreement exceeds agreement arising through random chance; this is valuable in situations where species diversity is poor and the virtual analyst could produce fairly high percent agreement by simply guessing the most common species every time. We calculated kappa using the irr package in R version 3.2.2 (R

Core Team, 2015). To assess agreement on tree identification, kappa was computed separately for genus and species. To assess size class agreement, trees were first binned into the following dbh classes commonly employed in forestry: 0-7.6 cm (0-3 in.), 7.6-15.2 cm (3-6 in.), 15.2-30.5 cm (6-12 in.), 30.5-45.7 cm (12-18 in.), 45.7-61.0 cm (18-24 in.), 61.0-76.2 cm (24-30 in.), and >76.2 cm (>30 in.). Then we calculated a weighted kappa on these ordinal data, wherein virtual dbh estimates were penalized more harshly as their distance from the observed size class increased.

In addition to kappa, we used linear regression analysis in R version 3.2.2 (R Core Team, 2015) to compare dbh estimates from the virtual survey to field measurements. In the regression equation, a slope of 1 would indicate that the set of virtual dbh estimates was proportional to field measurements, a slope <1 would indicate dbh was underestimated in the virtual survey, and a slope >1 would indicate dbh was overestimated in the virtual survey. The coefficient of determination (R^2) was used to evaluate the consistency of virtual dbh estimates, where a higher R^2 value indicated a more consistent relationship between virtual dbh estimates and field measurements, and a lower R^2 value indicated a weaker pattern between virtual estimates and field measurements. Regressions were computed for each municipality individually and for all three municipalities combined.

Results

We conducted both field and virtual surveys for 115 street segments across the three study municipalities (Table 1). The virtual survey of Reading and Wyoming took 11.2 hours for one analyst to complete (time was not recorded for Mt. Healthy). This was decidedly faster than the 28.5 hours needed for a two-person field team to collect the same tree variables in Reading and

Wyoming. Although we inventoried the same street segments in the two surveys, there was a temporal mismatch in data capture for every street segment. The field survey took place September to October 2013, while GSV imagery on the study street segments was acquired from July through September (leaf-on season) across the following years: 2009 (2% of imagery), 2011 (38%), 2014 (56%), 2015 (3%), and 2016 (<1%). We encountered 597 trees in the field survey and 606 trees in the GSV virtual survey (Table 1). Tree counts were the same between the two surveys for 67% of street segments. The average difference in tree counts between the field and virtual surveys was 0.07 trees per street segment, and the single largest discrepancy for a segment was seven trees. Over 93% of trees encountered in the field survey were matched between the two surveys based on their location, dbh, and genus (Table 1).

For matched trees, the genus identification was in agreement between the field and virtual surveys for 90% of trees (kappa = 0.88, $p < 0.001$), and agreement was higher for larger trees (Fig. 1). At the species level, the level of agreement for tree identification was 66% of trees (kappa = 0.64, $p < 0.001$). Agreement was very high for common species, particularly when that species was the only member of its genus represented in the study area. For example, *Pyrus calleryana* was the most common species encountered in the field (14% of matched trees), and the field and virtual surveys agreed for nearly 99% of individuals. Similarly, *Gleditsia triacanthos* and *Zelkova serrata* were both common (6% and 5% of matched trees, respectively) and identified with 100% agreement between the two surveys. On the other hand, agreement was much lower for genera with many species represented in the field survey. For instance, six *Ulmus* species represented 7% of the field survey, but there was 0% agreement at the species level between the two surveys. Likewise, agreement was low (37%) for the eight *Quercus* species comprising 7% of the field survey.

Virtual survey estimates of dbh placed 67% of matched trees into the same size bin (weighted kappa = 0.73, $p < 0.001$). Regression slopes were <1 in each municipality, indicating that the analyst underestimated dbh in the virtual survey (Fig. 2). However, the analyst improved his estimates over the course of the study by evaluating his performance after each municipality before moving on to the next, as indicated by increasing R^2 values and slopes increasing toward 1 as he progressed from Mt. Healthy to Reading to Wyoming (Fig. 2).

Discussion

GSV imagery shows promise as a tool for conducting virtual surveys of street trees. As expected, the virtual survey was conducted by a single person more quickly, at lower cost, and with reduced safety risks as compared to the field survey. The intended use of street tree data ultimately dictates what is considered acceptable data quality, so we hesitate to make judgments about the ability to generate usable data with this approach. In general, the virtual survey produced similar data to the field survey, which is promising for the continued implementation of GSV as a geospatial tool to aid research and practice in urban forestry. For example, the similar number of total trees observed in the field and virtual surveys falls within the range of error expected when conducting a sample inventory for an i-Tree Streets study (i-Tree, 2016), and most trees were observed in both surveys (Table 1). When trees were not identified in both surveys, it could be attributable to two main issues. First, GSV imagery for the study area was acquired from 2009-2016, but no imagery was acquired in 2013 when the field survey took place. Thus, a number of trees were planted or removed in the time between the two surveys. Second, in locations without a sidewalk to delineate the public right-of-way, the virtual survey

analyst occasionally had difficulty determining whether a tree should be counted as a street tree or not.

While our findings provide the first gauge of how well GSV can be used to conduct virtual street tree surveys, several limitations and recommendations should be considered. High agreement in genus identification suggests that analysts can use GSV to reliably identify most trees to the genus level, especially for larger trees (Fig.1). Larger trees have more leaves and larger boles with defining bark characteristics to aid identification, and they extend closer to the GSV panorama center. Small trees with very few leaves are difficult to see in GSV, especially if the image quality is marginal at that location. While genus identification was fairly reliable, species identification was less successful, particularly for genera like *Acer, Quercus,* and *Ulmus* that had multiple species represented in our sample. In our study, the analyst had never visited the study area, but a local urban forester with knowledge of the species planted within a municipality would have an advantage in identifying trees using GSV imagery. Providing analysts with training materials including photos and lists of distinguishing characteristics for local street trees could improve species identification. Note that some of the observed disagreement for species identification likely stems from the field crews and virtual analyst recognizing the same type of tree but assigning it two different names; for example, a hybrid Freeman maple may have been recorded as any one of *Acer x freemanii, A. rubrum,* or *A. saccharinum.*

The analyst's estimates of dbh improved markedly as he learned from his performance in one municipality and applied that insight to the next (Fig. 2). For future studies, we recommend developing a photographic guide showing trees of various diameters in GSV imagery so analysts can compare trees in their studies to reference trees of known dbh. While larger trees were more

readily identified at the genus level, it was more difficult to accurately estimate dbh for larger trees, as evidenced by less consistent estimates for larger trees (Fig. 2). As dbh estimates were generally more accurate for smaller trees, the relatively high proportion of smaller trees in Wyoming may partly explain the improved success estimating dbh in that community (Fig. 2). We used manual estimates of dbh because we wanted to conduct a free analysis without introducing advanced computing requirements, but it may be possible to use image measurement software to automate dbh estimation. This automation could improve consistency across multiple analysts and save time, and may thus be valuable even if software-based estimates of dbh are somewhat inaccurate. More broadly, we are aware of ongoing efforts to automate the use of GSV imagery to generate a suite of tree data including tree counts, health, size, and species, but these techniques are still under development and require advanced computer skills beyond the reach of most communities. The manual approach proposed here provides a lower-tech alternative that is widely accessible to those with limited technological capabilities.

We used GSV to replicate an existing field survey for the sake of comparison, but practical applications could include updating outdated street tree inventories and generating new street tree data. In such cases, groundtruthing should be performed to assess the quality of tree data generated using GSV. We worked in lower density urban areas, so we do not know how well our approach would work in denser urban centers where street trees may be obscured by traffic or infrastructure in GSV imagery. This analysis was conducted by a single individual who has a background in field botany. As such, we cannot estimate inter-operator consistency or describe how reliably this approach could be implemented by volunteers or other non-experts. It is possible that our approach could be used by an urban forester to validate field data collected by volunteers.

Virtual street tree surveys in GSV are appealing for several reasons. They can be done with less time, labor, money, and safety risks than field surveys. Virtual surveys can be conducted year-round regardless of season or weather conditions. As opposed to other geospatial technologies used in urban forestry, GSV is freely available and easy to use. In our comparison of virtual survey data to field-based data, we found high levels of agreement in tree counts and genus identification, and identification was easiest for large trees. Identifying trees to the species level was more difficult. The analyst's ability to manually estimate dbh improved markedly over time, demonstrating a learning curve that could be eliminated in the future by implementing image measurement software. In light of these results, GSV shows strong potential for complementing field data collection as a means of generating information about street trees.

References

Alonzo, M., Bookhagen, B., Roberts, D.A., 2014. Urban tree species mapping using hyperspectral and lidar data fusion. Remote Sensing of Environment 148, 70-83.

Alonzo, M., McFadden, J.P., Nowak, D.J., Roberts, D.A., 2016. Mapping urban forest structure and function using hyperspectral imagery and lidar data. Urban Forestry & Urban Greening 17, 135-147.

Anguelov, D., Dulong, C., Filip, D., Frueh, C., Lafon, S., Lyon, R., Ogale, A., Vincent, L., Weaver, J., 2010. Google Street View: capturing the world at street level. Computer 6, 32-38.

Badland, H.M., Opit, S., Witten, K., Kearns, R.A., Mavoa, S., 2010. Can virtual streetscape audits reliably replace physical streetscape audits? Journal of Urban Health 87, 1007-1016.

Berland, A., Hopton, M.E., 2014. Comparing street tree assemblages and associated stormwater benefits among communities in metropolitan Cincinnati, Ohio, USA. Urban Forestry & Urban Greening 13, 734-741.

Berland, A., Schwarz, K., Herrmann, D.L., Hopton, M.E., 2015. How environmental justice patterns are shaped by place: terrain and tree canopy in Cincinnati, Ohio, USA. Cities and the Environment (CATE) 8, Article 1.

Clarke, P., Ailshire, J., Melendez, R., Bader, M., Morenoff, J., 2010. Using Google Earth to conduct a neighborhood audit: reliability of a virtual audit instrument. Health & Place 16, 1224-1229.

ESRI, 2014. ArcGIS for Desktop version 10.3. Redlands, CA.

i-Tree, 2016. i-Tree Streets. http://www.itreetools.org/streets/index.php (accessed August 2016).

Landry, S.M., Chakraborty, J., 2009. Street trees and equity: evaluating the spatial distribution of an urban amenity. Environment and Planning A 41, 2651-2670.

Li, X., Zhang, C., Li, W., Kuzovkina, Y.A., Weiner, D., 2015a. Who lives in greener neighborhoods? The distribution of street greenery and its association with residents' socioeconomic conditions in Hartford, Connecticut, USA. Urban Forestry & Urban Greening 14, 751-759.

Li, X., Zhang, C., Li, W., Ricard, R., Meng, Q., Zhang, W., 2015b. Assessing street-level urban greenery using Google Street View and a modified green view index. Urban Forestry & Urban Greening 14, 675-685.

McGee, J.A., III, Day, S.D., Wynne, R.H., White, M.B., 2012. Using geospatial tools to assess the urban tree canopy: decision support for local governments. Journal of Forestry 110, 275-286.

McPherson, E.G., van Doorn, N., de Goede, J., 2016. Structure, function and value of street trees in California, USA. Urban Forestry & Urban Greening 17, 104-115.

Mullaney, J., Lucke, T., Trueman, S.J., 2015. A review of benefits and challenges in growing street trees in paved urban environments. Landscape and Urban Planning 134, 157-166.

O'Neil-Dunne, J., MacFaden, S., Royar, A., 2014. A versatile, production-oriented approach to high-resolution tree-canopy mapping in urban and suburban landscapes using GEOBIA and data fusion. Remote Sensing 6, 12837-12865.

Pedlar, J.H., McKenney, D.W., Allen, D., Lawrence, K., Lawrence, G., Campbell, K., 2013. A street tree survey for Canadian communities: protocol and early results. The Forestry Chronicle 89, 753-758.

R Core Team, 2015. R: a language and environment for statistical computing. https://www.R-project.org (accessed August 2016).

Rousselet, J., Imbert, C.-E., Dekri, A., Garcia, J., Goussard, F., Vincent, B., Denux, O., Robinet, C., Dorkeld, F., Roques, A., Rossi, J.-P., 2013. Assessing species distribution using Google Street View: a pilot study with the pine processionary moth. PLoS ONE 8, e74918.

San Souci, J., Hanou, I., Puchalski, D., 2009. High-resolution remote sensing image analysis for early detection and response planning for emerald ash borer. Photogrammetric Engineering & Remote Sensing 75, 905-909.

US Census Bureau, 2016. American FactFinder. http://factfinder.census.gov/ (accessed August 2016).

Ward, K.T., Johnson, G.R., 2007. Geospatial methods provide timely and comprehensive urban forest information. Urban Forestry & Urban Greening 6, 15-22.

Figure legends

Figure 1. Agreement between the field and virtual surveys in terms of tree identification at the genus and species levels. Dbh classes are: 1 (0-7.6 cm), 2 (7.6-15.2 cm), 3 (15.2-30.5 cm), 4 (30.5-45.7 cm), 5 (45.7-61.0 cm), 6 (61.0-76.2 cm), and 7 (>76.2 cm). White circles indicate the number of trees in each dbh class.

Figure 2. Linear regressions assessing the relationship between field measurements and virtual survey estimates of dbh. The regression (solid) and 1-to-1 (dashed) lines are shown for reference. The analyst completed estimates in the order of Mt. Healthy, Reading, and then Wyoming, and evaluated his performance in each municipality before moving on to the next.

Figure 1.

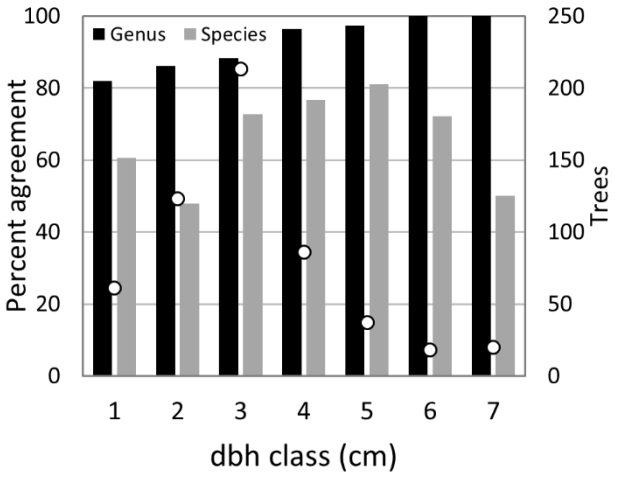
Genus
Species
Percent agreement
Trees
dbh class (cm)
100
80
60
40
20
0
250
200
150
100
50
0
1
2
3
4
5
6
7

Figure 2.

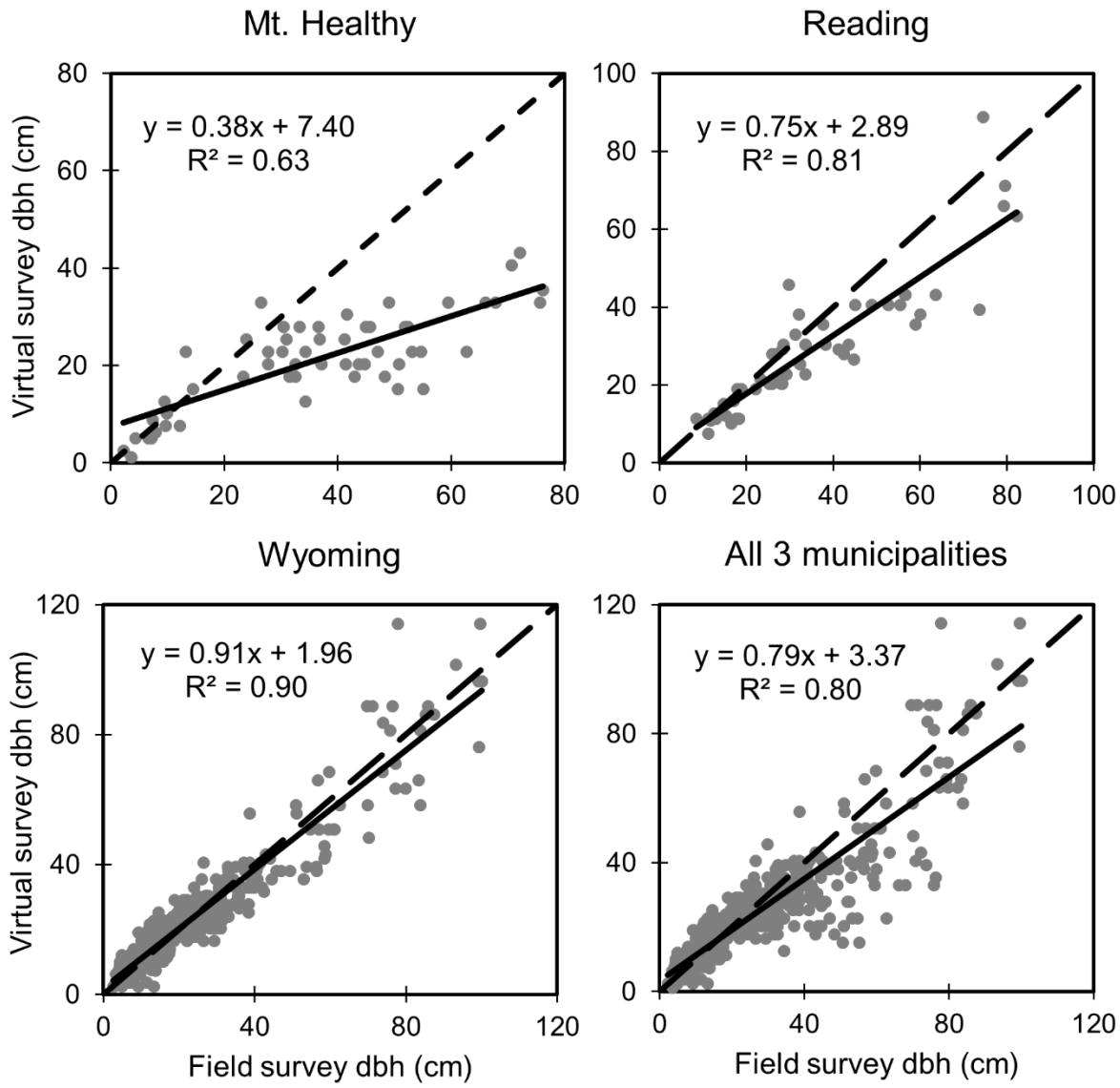
Mt. Healthy
y = 0.38x + 7.40
R² = 0.63
Reading
y = 0.75x + 2.89
R² = 0.81
Wyoming
y = 0.91x + 1.96
R² = 0.90
All 3 municipalities
y = 0.79x + 3.37
R² = 0.80
Virtual survey dbh (cm)
Field survey dbh (cm)

Table 1. Summary of trees inventoried in the field and virtual surveys, and number of trees matched between the two surveys based on their geographic location, species, and dbh

Municipality	Street segments (total km)	Trees (field)	Trees (virtual)	Matched trees (% of field total)
Mt. Healthy	31 (4.4 km)	63	69	56 (89%)
Reading	49 (5.8 km)	64	55	54 (84%)
Wyoming	35 (5.7 km)	470	482	448 (95%)
All 3 combined	115 (15.9 km)	597	606	558 (94%)

APPENDIX 2: TREE ESTIMATION REFERENCE SHEET

APPENDIX 3: LiDAR DEPRESSION FINDER – FULL PYTHON SCRIPT

```
1.  import arcpy
2.  import os.path
3.  arcpy.env.overwriteOutput = True
4.
5.  #Variables
6.
7.  ##File Locations
8.
9.  LAS_File = arcpy.GetParameterAsText(0)
10. LASD_Name = os.path.basename(LAS_File.rstrip(".las"))
11. SC = "SC_" + LASD_Name
12.
13. Out_Loc = arcpy.GetParameterAsText(1)
14. if arcpy.Exists(Out_Loc) == False:
15.     arcpy.CreateFolder_management(os.path.dirname(Out_Loc), os.path.basename(Out_Loc))
16. Out_Loc = Out_Loc + "\\"
17. Out_GDB = Out_Loc + "DansLDF_Outputs.gdb"
18. if arcpy.Exists(Out_GDB) == False:
19.     arcpy.CreateFileGDB_management(Out_Loc, os.path.basename(Out_GDB.rstrip(".gdb")))
20.
21. arcpy.env.workspace = Out_GDB
22.
23. ##Numerical Values
24.
25. CI = arcpy.GetParameter(2)
26. TH = arcpy.GetParameter(3)
27.
28. ##Clean Up Options
29.
30. Del_LASD = arcpy.GetParameter(4)
31. Del_SC = arcpy.GetParameter(5)
32. Del_SC_Poly = arcpy.GetParameter(6)
33. Del_SC_Polyj = arcpy.GetParameter(7)
34. Del_NBRS = arcpy.GetParameter(8)
35. Del_STAT = arcpy.GetParameter(9)
36. Del_PD = arcpy.GetParameter(10)
37.
38. #Create LAS Dataset
39.
40. arcpy.CreateLasDataset_management(LAS_File, Out_Loc + LASD_Name, "NO_RECURSION", "", "", "COMP
      UTE_STATS", "", "NO_FILES")
41. arcpy.AddMessage('LAS Dataset {0} Created'.format(LASD_Name))
42.
43. #Create LAS Dataset Layer
44. ##This is to filter out only bare earth points for contour creation
45.
46. arcpy.MakeLasDatasetLayer_management(Out_Loc + LASD_Name + ".lasd", LASD_Name + "_lyr", 2, "", "E
      XCLUDE_UNFLAGGED", "", "", "", "", )
47. arcpy.AddMessage("LAS Dataset Layer {0} Created".format(LASD_Name + "_lyr"))
48.
```

```
49.  #Create Contours from LAS Dataset Layer
50.
51.  arcpy.CheckOutExtension("3D")
52.  arcpy.SurfaceContour_3d(LASD_Name + "_lyr", SC, CI, "", "", 2, "", "", "", "")
53.  arcpy.CheckInExtension("3D")
54.  arcpy.AddMessage("Surface Contour {0} Created".format(SC))
55.
56.  #Extract Looped Contour Lines from Surface Contour
57.  ##This is done by comparing line geopmetry; contour lines that form closed loops will have the same start
       and end points
58.  ##By extracting only those loops above a certain threshold length, many artifacts from contour productio
       n can be eliminated
59.
60.  ##Add geometric attribute fields to attribute table
61.
62.  arcpy.AddGeometryAttributes_management(SC, "LENGTH;LINE_START_MID_END", "METERS")
63.  arcpy.AddMessage("Geometric Attributes Added to {0}".format(SC))
64.
65.  ##create and calculate loop identification field
66.
67.  arcpy.AddField_management(SC, "XY_Diff", "DOUBLE")
68.  arcpy.CalculateField_management(SC, "XY_Diff", "Abs((START_X + START_Y) - (END_X + END_Y))")
69.  arcpy.AddMessage("XY_Diff Field Added and Calculated")
70.
71.  ##Identify and extract loops of significant size
72.
73.  arcpy.MakeFeatureLayer_management(SC, SC + "_lyr", "XY_Diff = 0 AND LENGTH >= {0}".format(TH))
74.
75.  ##Verify Loops of Significant Size
76.
77.  L_Num = arcpy.GetCount_management(SC + "_lyr")
78.  if L_Num == 0:
79.     arcpy.AddWarning("No Loops Found")
80.  else:
81.     arcpy.AddMessage("{0} Loops Longer than {1} Meters Found".format(L_Num, TH))
82.
83.  #Convert Loops to Polygons
84.
85.  arcpy.FeatureToPolygon_management(SC + "_lyr", SC + "_Poly")
86.  arcpy.AddMessage("Converted Loops to Polygons: {0}".format(SC + "_Poly"))
87.
88.  #Spatially Join Contour Line Values to Blank Polygons
89.  ##The Feature To Polygon tool does not carry over line attributes like it should so they have to be joined i
       n from the contour lines
90.  ##All of the polygons except for the centers of concentric hills or depressions will touch multiple contour l
       ines
91.  ##The spatial join will attempt to account for this by including several contour statistics
92.
93.  ##Field Mappings
94.
95.  FMS = arcpy.FieldMappings()
96.
97.  FM_CF = arcpy.FieldMap() ##First contour of the intersection
```

```
98.  FM_CL = arcpy.FieldMap() ##Last contour of the intersection
99.  FM_CC = arcpy.FieldMap() ##Number of contours intersected
100. FM_LF = arcpy.FieldMap() ##Length of the first contour line
101. FM_LL = arcpy.FieldMap() ##Length of the last contour line
102.
103. FM_CF.addInputField(SC + "_lyr", "Contour")
104. FM_CL.addInputField(SC + "_lyr", "Contour")
105. FM_CC.addInputField(SC + "_lyr", "Contour")
106. FM_LF.addInputField(SC + "_lyr", "Shape_Length")
107. FM_LL.addInputField(SC + "_lyr", "Shape_Length")
108.
109. FM_CF.mergeRule = "First"
110. FM_CL.mergeRule = "Last"
111. FM_CC.mergeRule = "Count"
112. FM_LF.mergeRule = "First"
113. FM_LL.mergeRule = "Last"
114.
115. OF_CF = FM_CF.outputField
116. OF_CL = FM_CL.outputField
117. OF_CC = FM_CC.outputField
118. OF_LF = FM_LF.outputField
119. OF_LL = FM_LL.outputField
120.
121. OF_CF.name = "CFirst"
122. OF_CL.name = "CLast"
123. OF_CC.name = "CCount"
124. OF_LF.name = "LFirst"
125. OF_LL.name = "LLast"
126.
127. FM_CF.outputField = OF_CF
128. FM_CL.outputField = OF_CL
129. FM_CC.outputField = OF_CC
130. FM_LF.outputField = OF_LF
131. FM_LL.outputField = OF_LL
132.
133. FMS.addFieldMap(FM_CF)
134. FMS.addFieldMap(FM_CL)
135. FMS.addFieldMap(FM_CC)
136. FMS.addFieldMap(FM_LF)
137. FMS.addFieldMap(FM_LL)
138.
139. ##Spatial Join
140.
141. arcpy.SpatialJoin_analysis(SC + "_Poly", SC + "_lyr", SC + "_Polyj", "", "", FMS)
142. arcpy.AddMessage("Joined Contour Lines to Polygons, {0} Created".format(SC + "_Polyj"))
143.
144. #Polygon Neighbors
145. ##Comparison between neighbor vallues will be the basis for depression location
146.
147. arcpy.PolygonNeighbors_analysis(SC + "_Polyj", LASD_Name + "_NBRS", ["OBJECTID", "CFirst", "CLast", "C
       Count", "LFirst", "LLast", "Shape_Length"])
148. arcpy.AddMessage("Polygon Neighbors Table {0} Created".format(LASD_Name + "_NBRS"))
149.
```

```
150.  #Summarize Polygon Neighbors and Rejoin to Polygon Contours
151.  ##The Polygon Neighbor tool generates every possible relationhip, this will sumarize the relationships so t
        hat each polygon has only one set of neighbor values
152.
153.  arcpy.Statistics_analysis(LASD_Name + "_NBRS", LASD_Name + "_STAT", [["src_CFirst", "MEAN"], ["nbr_C
        First", "MIN"], ["nbr_CLast", "MIN"], ["nbr_CCount", "MEAN"], ["src_LFirst", "MEAN"], ["src_Shape_Len
        gth", "MEAN"]], "src_OBJECTID")
154.  arcpy.AddMessage("Statistics Table {0} Created".format(LASD_Name + "_STAT"))
155.
156.  arcpy.JoinField_management(SC + "_Poly", "OBJECTID", LASD_Name + "_STAT", "src_OBJECTID")
157.  arcpy.AddMessage("Statistics Table Joined To {0}".format(SC + "_Poly"))
158.
159.  #Extract Depressions
160.  ##This is done using a long where clause containing the following parts
161.  ##Frequency = 1 - Only loops with one neighbor like hill tops and depression centers will be considered
162.  ##MEAN_src_CFirst = MIN_nbr_CFirst - Limits the selection to only loops that are the lowest contoured lo
        op that their neigbor is touching
163.  ##MIN_nbr_CFirst < MIN_nbr_CLast - Verifies that the neighbors indide edge is a smaller contour than the
         outside edge
164.  ##MEAN_nbr_CCount = 2 - Verifies that the neighbor is part of a concentric ring arrangement, reducing er
        ror from odd relationships
165.  ##MEAN_src_LFirst = MEAN_src_Shape_Length - Eliminates a significant number, but not all, of self inters
        ecting line artifacts
166.
167.  arcpy.MakeFeatureLayer_management(SC + "_Poly", SC + "_DEPlyr", "FREQUENCY = 1 AND MEAN_src_CF
        irst = MIN_nbr_CFirst AND MIN_nbr_CFirst < MIN_nbr_CLast AND MEAN_nbr_CCount = 2 AND MEAN_s
        rc_LFirst = MEAN_src_Shape_Length")
168.
169.  DEP_Count = arcpy.GetCount_management(SC + "_Deplyr")
170.  DEP_Count = int(DEP_Count.getOutput(0))
171.  if DEP_Count == 0:
172.      arcpy.AddWarning("No Potential Depressions Found With Given Parameters")
173.  else:
174.      arcpy.AddMessage("{0} Potential Depressions Identified".format(DEP_Count))
175.
176.  arcpy.CopyFeatures_management(SC + "_DEPlyr", LASD_Name + "_PD")
177.  arcpy.AddMessage("Potential Depressions Extracted to {0}".format(LASD_Name + "_PD"))
178.
179.  #Clean Up
180.  ##Optional deletion of unwanted output products
181.
182.  if Del_LASD == True:
183.      arcpy.Delete_management(Out_Loc + LASD_Name + ".lasd")
184.      arcpy.AddMessage(LASD_Name + ".lasd Deleted")
185.  if Del_SC == True:
186.      arcpy.Delete_management(SC)
187.      arcpy.AddMessage(SC + " Deleted")
188.  if Del_SC_Poly == True:
189.      arcpy.Delete_management(SC + "_Poly")
190.      arcpy.AddMessage(SC + "_Poly Deleted")
191.  if Del_SC_Polyj == True:
192.      arcpy.Delete_management(SC + "_Polyj")
193.      arcpy.AddMessage(SC + "_Polyj Deleted")
```

```
194.  if Del_NBRS == True:
195.      arcpy.Delete_management(LASD_Name + "_NBRS")
196.      arcpy.AddMessage(LASD_Name + "_NBRS Deleted")
197.  if Del_STAT == True:
198.      arcpy.Delete_management(LASD_Name + "_STAT")
199.      arcpy.AddMessage(LASD_Name + "_STAT Deleted")
200.  if Del_PD == True:
201.      if DEP_Count == 0:
202.          arcpy.Delete_management(LASD_Name + "_PD")
203.          arcpy.AddMessage(LASD_Name + "_PD Deleted")
```

(Script formatted for MS Word using tool at: http://www.planetb.ca/syntax-highlight-word)